VOLUME 30 JANUARY 1952 NUMBER 1

Canadian Journal of Physics

Editor: G. M. VOLKOFF

Published by THE NATIONAL RESEARCH COUNCIL
OTTAWA CANADA

CANADIAN JOURNAL OF PHYSICS

(Formerly Section A, Canadian Journal of Research)

The CANADIAN JOURNAL OF PHYSICS is published bimonthly by the National Research Council of Canada under the authority of the Chairman of the Committee of the Privy Council on Scientific and Industrial Research. Matters of general policy are the responsibility of a joint Editorial Board consisting of members of the National Research Council of Canada and the Royal Society of Canada.

The National Research Council of Canada publishes also: *Canadian Journal of Botany, Canadian Journal of Chemistry, Canadian Journal of Medical Sciences, Canadian Journal of Technology, Canadian Journal of Zoology.*

Manuscripts should be addressed to:

DR. LÉO MARION,
Editor-in-Chief,
Canadian Journal of Physics,
National Research Council,
Ottawa, Canada.

Each manuscript should be typewritten, double-spaced, and the original and one extra copy submitted (see **Notice to Contributors,** inside of back cover).

Subscription, renewals, and orders for back numbers should be addressed to:

Administrative Services,
National Research Council,
Ottawa, Canada.

Subscription rate: $3.00 a year; single numbers: 50 cents. Special rates can be obtained for subscriptions to more than one of the Journals published by the National Research Council.

*Deceased.

VOLUME 30 1952

Canadian Journal of Physics

Published by THE NATIONAL RESEARCH COUNCIL
OTTAWA CANADA

Canadian Journal of Physics

Issued by THE NATIONAL RESEARCH COUNCIL OF CANADA

VOLUME 30 JANUARY, 1952 NUMBER 1

A LINEAR RADIO-FREQUENCY MASS SPECTROMETER[1]

BY P. A. REDHEAD

ABSTRACT

A radio-frequency mass spectrometer is described using a system of electrodes similar to a linear accelerator to obtain velocity selection of tin ions. A beam of ions is directed through a series of axial radio-frequency fields with alternate polarity, and the ion with a velocity (proportional to $M^{-\frac{1}{2}}$) such that it traverses the individual stages in approximately one-half period of the radio-frequency field, acquires maximum energy from this field. A retarding field region prevents all ions, except the ion with maximum energy, from reaching the collector. An ion beam of large cross section is used. With 20 radio-frequency stages a mass resolution of 1% is obtained.

INTRODUCTION

All mass spectrometers employing magnetic deflection to obtain mass selection require accurately aligned apertures which have the disadvantage of limiting the useful ion current. Numerous proposals have been put forward for using time-varying electric fields to obtain velocity selection in a beam of ions. In these systems the ion current is not so severely limited by apertures. The use of transverse deflecting fields has been suggested by Smythe (11), Mattauch (12), and Hintersberger (7). Time-of-flight methods using a short pulse of ions moving linearly have been suggested by Keller (8), Stephens (13), and Cameron (3). The use of a pulsed ion beam rotating in a magnetic field has been described by Goudsmit (4), and Smith (10). The use of axial fields in a linear system has been described by Bennett (1), Boyd (2), Weisz (14), and Heil (5). This last method has the advantage of in no way limiting the cross section of the ion beam by apertures or deflecting electrodes and leads to an instrument with a high ion current and medium resolution.

The instrument herein described resembles a linear accelerator similar to the type developed by Sloan and Coates (9), but without drift spaces, and where the time of transit through the individual radio-frequency fields is comparable to the period of the applied field. An ion with the synchronous velocity, determined by its e/m ratio and the initial d-c. accelerating potential, traverses each of the individual fields in one-half period. This synchronous ion acquires more energy from the radio-frequency fields than ions with any other velocity and is thus able to pass through a d-c. retarding field and reach the collector. The use of transit times comparable with the period of the applied field causes rejection of ions with harmonic masses (ions whose masses are integral

[1] *Manuscript received August 10, 1951.*
Contribution from the Radio and Electrical Engineering Division, National Research Council, Ottawa, Canada. Issued as N.R.C. No. 2636.

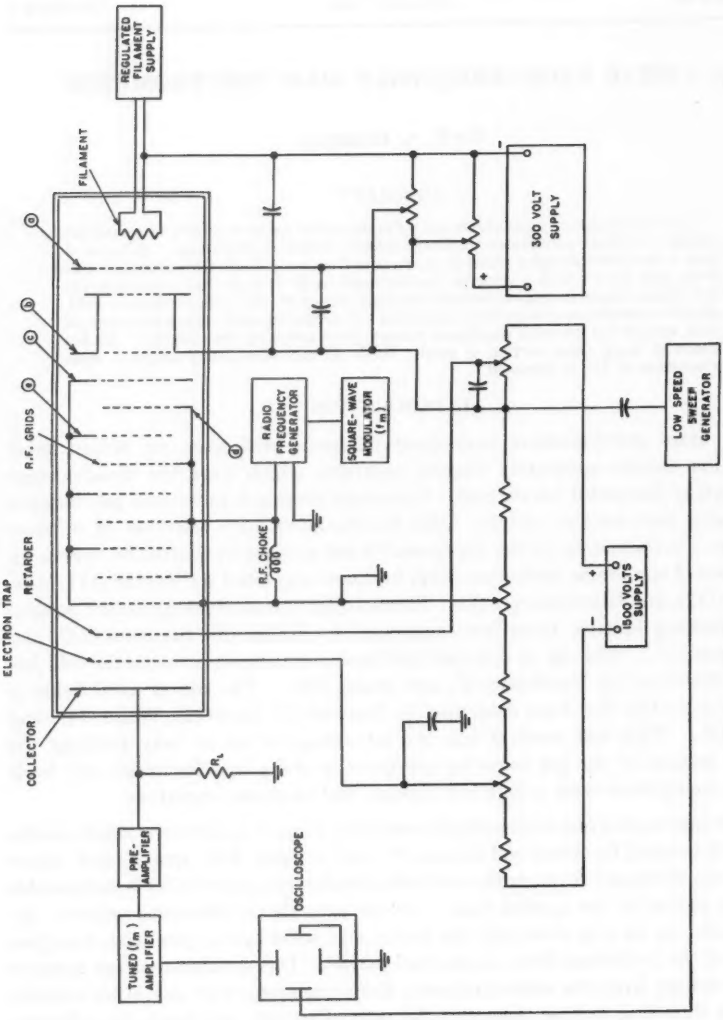

FIG. 1. Schematic diagram of tube and auxiliary circuits.

multiples of the mass of the synchronous ion). Many other time-of-flight devices do not discriminate against harmonic masses.

OPERATION OF THE TUBE

The tube and circuits are shown schematically in Fig. 1. Electrons from the filament are accelerated by grid (a) into the field-free cylinder (b) where ions are produced by collision. The positive ions are accelerated out of cylinder (b) toward a grid (c) by a d-c. potential (V_0). The ions then pass through a series of equally spaced grids which are all at d-c. ground. Alternate grids are connected to a radio-frequency generator of pulsatance ω. The electric field between (c) and (d) is

$$E \sin (\omega t + \theta)$$

and between (d) and (e) is

$$-E \sin (\omega t + \theta).$$

A positive ion entering the space c–d just as the field is going negative will be accelerated; moreover, if the transit angle of the ion from (c) to (d) is π radians, it will enter the second gap at the correct phase to be accelerated once more. This process is repeated through all the gaps down the tube. If the energy gained by the ion is large it will, of course, slip out of phase with the fields; however, this effect is not troublesome in practice.

An exact solution of the equations of the ionic motion is only possible by graphical methods. An approximate analytical solution may be obtained by assuming small signal conditions. ΔW, the energy gained by an ion in traversing N gaps, may be derived in terms of α, the unmodulated transit angle for one gap (see Appendix I).

$$\frac{\Delta W}{\epsilon V} = \frac{1}{\alpha}\left\{2 \cos \frac{N}{2}\alpha - 4 \cos \left(\frac{N}{2} - 1\right)\alpha + 4 \cos \left(\frac{N}{2} - 2\right)\alpha \ldots - 4 \cos \alpha + 2\right\}$$

where V is the peak radio-frequency voltage across each gap. The quantity, $\Delta W/\epsilon V$, is plotted as a function of α in Fig. 2 for $N = 8$, and in Fig. 3 for

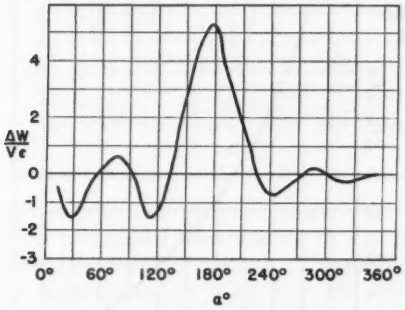

FIG. 2. Energy gained by an ion as a function of the unmodulated gap transit angle for $N = 8$.

$N = 20$. Comparison of the two graphs shows clearly that as N is increased the main peak becomes narrower and higher and the value of α at this peak

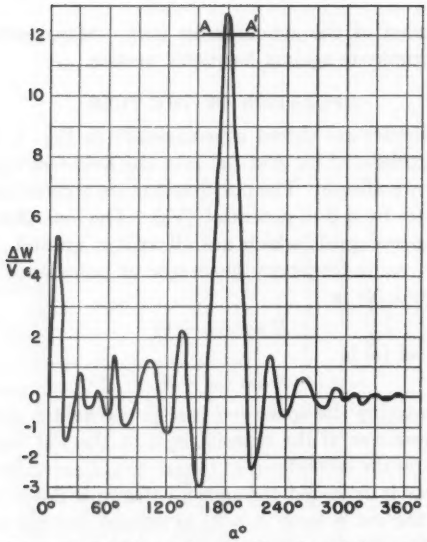

FIG. 3. Energy gained by an ion as a function of the unmodulated gap transit angle for $N = 20$.

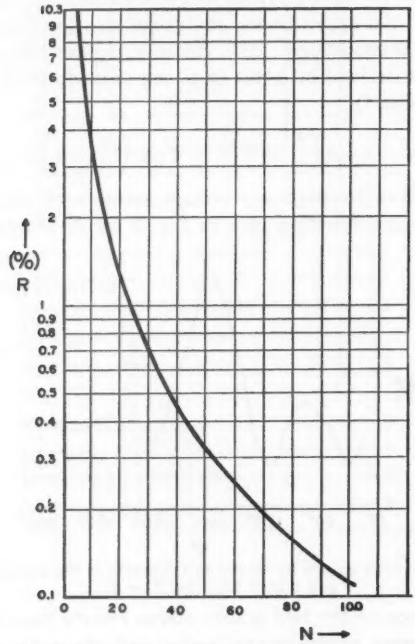

FIG. 4. Variation of percentage resolution as a function of N.

(α_0) approaches π, and thus the optimum entrance phase angle (θ_0) approaches zero. (See Equation 4(a), Appendix I.)

The improvement in resolution obtained by increasing the number of gaps N, is evident in the narrowing of the main peak. The resolution has been computed (see Appendix II) as a function of N and is shown in Fig. 4. These curves are for

$$\frac{\delta V}{V} = \frac{1}{12},$$

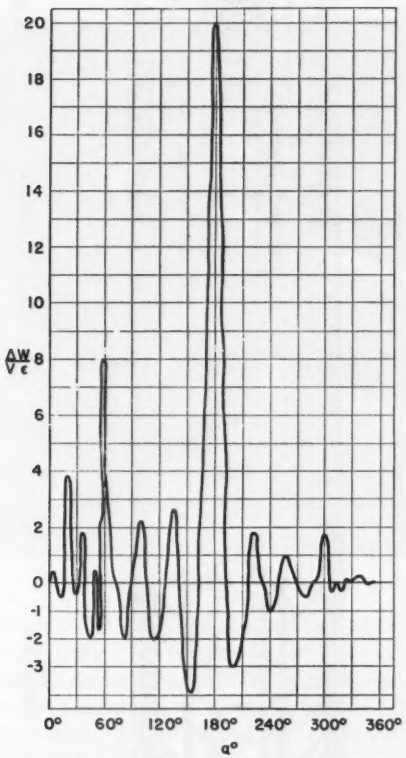

FIG. 5. Energy gained as a function of α for $N = 20$, using square waves.

where δV is the effective potential difference between the height of the main peak and the retarding potential (corresponding to AA' in Fig. 3), and V is the peak value of the radio-frequency voltage. This simple analysis predicted a mass resolution of approximately 1% for $N = 20$, which was achieved in practice.

The use of square waves instead of a sinusoidal wave form will sensibly improve the resolution (6). Fig. 5 shows the quantity $\Delta W/\epsilon V$ as a function of α for $N = 20$, when the accelerator section is driven by a square wave.

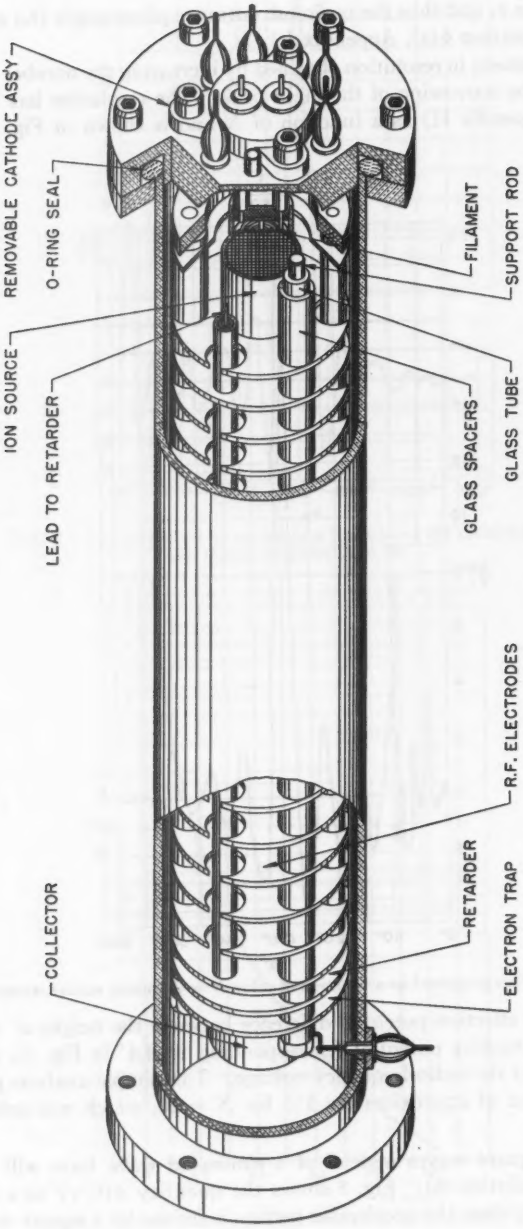

FIG. 6. Cutaway drawing of 20-stage tube.

DESCRIPTION OF THE TUBE AND CIRCUITS

The ion source is designed to produce an ion beam of large cross section. A flat spiralled filament emits electrons which are accelerated by the grid (*a*) into the cylinder (*b*); gas is introduced into this cylinder and ions are formed by collision. The negative field between (*b*) and (*c*) prevents the electrons from leaving the ion cylinder. The grid (*a*) is slightly positive with respect to the cylinder (*b*) to prevent positive ions from moving back towards the cathode.

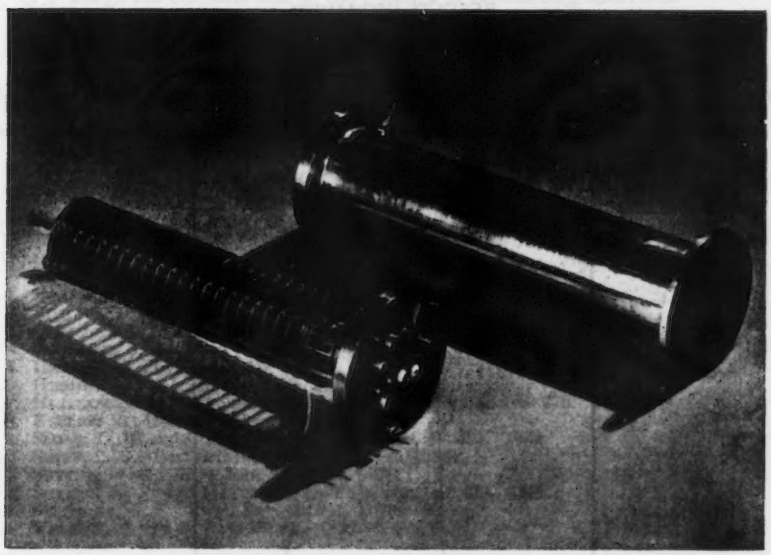

FIG. 7. Photograph of 20-stage tube.

The general construction of one model of the tube is shown in Fig. 6. Fig. 7 is a photograph of this tube. The electrode structure is mounted on three support rods and the electrodes are separated by glass spacers. The cathode assembly is mounted on a separate "0" ring seal for ease of replacement. The apertures in the electrodes (5/8 in. diameter) are covered with a knitted tungsten mesh of 95% transparency.

The accelerating stack of 21 electrodes is followed by the retarder grid and a grid to suppress secondary electrons from the collector cup as well as to prevent electrons released anywhere within the tube from reaching the collector. A preamplifier using a subminiature tube is connected directly to the collector lead-through. It is essential to keep the collector-to-ground capacitance small in order to reduce the time-constant of the electrometer grid circuit.

The circuit arrangement is shown schematically in Fig. 1. The radio-frequency oscillator is tunable over the range 3-13 Mc./sec., which, with the

radio-frequency gap spacing (1/4 in.) and the range of d-c. accelerating potentials used (200-600 volts), allows investigation of the mass range from 1 to 100. The radio-frequency oscillator is square-wave modulated at 400 cycles per second so that an a-c. system may be used to amplify the ion current. By electronically sweeping the initial d-c. accelerating voltage at a low frequency (ca. 1 cycle per second), the mass spectrum may be presented on a long-persistence cathode-ray tube as indicated in Fig. 1. When using a voltage sweep the mass scale is linear.

RESULTS OBTAINED

The experimental resolution obtained may be judged from Fig. 8, which is a tracing from the cathode-ray tube screen. Resolution may be sacrified for increased sensitivity, or vice versa, by adjustment of the retarder potential.

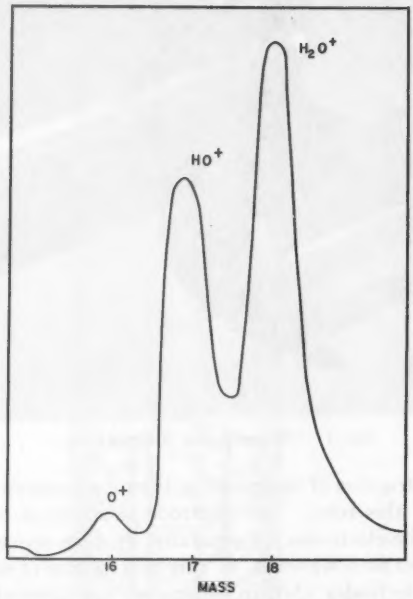

FIG. 8. Mass spectrum of water vapor traced from cathode-ray tube screen.

Both metal and glass envelope tubes have been operated, either with d-c. amplifiers and graphic recorders, or with the previously described dynamic presentation system. The tube has been found to operate satisfactorily from a pressure of 10^{-6} mm. to 10^{-4} mm.

The effect of the ions gaining velocity and slipping out of synchronism with the radio-frequency field has been investigated by applying d-c. setback voltages to the accelerating electrodes from an internal potentiometer system. This d-c. voltage will nullify the increase in velocity of the ions caused by the radio-

frequency fields. No change in sensitivity or resolution was observed with the setback voltages connected, even when the maximum radio-frequency voltage was used (15 volts r.m.s.). With the setback voltages it was found that the instrument was more nearly absolute in its measurement of ion masses.

COMMENTS

The advantages of this type of mass spectrometer over magnetic deflection type of equivalent resolution may be enumerated as:

(1). Large ion currents due to the absence of collimating apertures.
(2). Simplicity of construction; no accurate alignment is necessary.
(3). Absence of any magnet.
(4). Use of a-c. coupled amplifiers.

The main disadvantage is the need for calibration of the mass scale. Owing to the inadequacy of the analysis of the ion motion, the instrument is not absolute.

ACKNOWLEDGMENT

The writer wishes to acknowledge the valuable contributions made by Mr. G. A. Harrower (now at McGill University) to the success of this work.

REFERENCES

1. BENNETT, W. H. J. Applied Phys. 21: 143. 1950.
2. BOYD, R. L. F. Nature, 165: 142. 1950.
3. CAMERON, A. E. and EGGERS, D. F. Rev. Sci. Instruments, 19: 605. 1948.
4. GOUDSMIT, S. A. Phys. Rev. 74: 622. 1948.
5. HEIL, H. Z. Physik, 128: 36. 1950.
6. HENSON, A. F. J. Applied Phys. 21: 1063. 1950.
7. HINTERSBERGER, H. VON and MATTAUCH, J. Z. Physik, 106: 279. 1937.
8. KELLER, R. Helv. Phys. Acta, 22: 387. 1949.
9. SLOAN, D. H. and COATES, W. M. Phys. Rev. 46: 539. 1934.
10. SMITH, L. G. Rev. Sci. Instruments, 22: 115. 1951.
11. SMYTHE, W. R. Phys. Rev. 28: 1275. 1926.
12. SMYTHE, W. R. and MATTAUCH, J. Phys. Rev. 40: 429. 1932.
13. STEPHENS, W. E. Phys. Rev. 69: 691. 1946.
14. WEISZ, P. B. Phys. Rev. 70: 91. 1946.

(For Appendixes, see following pages.)

Appendix I

Let the electric field in the interaction gaps be represented by:

$$Ef(\omega t + \theta)$$

where $f(\omega t + \theta)$ represents the wave form of the applied voltage.

Then the force acting on an ion in the odd-numbered gaps is given by:

$$F = \frac{\epsilon V}{s} f(\omega t + \theta) \tag{1}$$

and in the even-numbered gaps by:

$$F = -\frac{\epsilon V}{s} f(\omega t + \theta)$$

where $\omega/2\pi$ is the frequency of the applied wave form,

θ is the phase of the applied wave form as the ion enters the first gap ($t = 0$),

s is the gap width.

Assuming that the unperturbed ion velocity v is large compared with any velocity changes produced by the accelerating wave form, then the energy acquired by the ion from the field is given by:

$$\Delta W = \Delta\left(\frac{1}{2}mv^2\right) = v\Delta(mv) = v\int F dt. \tag{2}$$

With N interaction gaps, where N is even, Equation (2) becomes:

$$\Delta W = \frac{v\epsilon V}{s}\left\{\int_0^{s/v} f(\omega t + \theta)dt - \int_{s/v}^{2s/v} f(\omega t + \theta)dt + \ldots - \int_{(N-1)s/v}^{Ns/v} f(\omega t + \theta)dt\right\}. \tag{3}$$

Integrating with respect to t

$$\Delta W = \frac{\epsilon V}{a}\left\{\Big[g(\omega t + \theta)\Big]_0^{s/v} - \Big[g(\omega t + \theta)\Big]_{s/v}^{2s/v} \ldots - \Big[g(\omega t + \theta)\Big]_{(N-1)s/v}^{Ns/v}\right\} \tag{4}$$

where $g(\omega t) = \int f(\omega t)\ d(\omega t)$, and $a = s\omega/v$ is the unmodulated interelectrode transit angle.

Maximizing ΔW with respect to variations in the entrance phase angle θ gives, for one solution

$$\frac{N}{2}a + \theta = \frac{N}{2}\pi. \tag{4a}$$

Inserting this condition for θ in Equation (4), one obtains

$$\frac{\Delta W}{\epsilon V} = \frac{1}{a}\left\{ - g\left[\frac{N}{2}\pi - \frac{N}{2}a\right] + 2g\left[\frac{N}{2}\pi - \left(\frac{N}{2} - 1\right)a\right]\right.$$
$$\left. - 2g\left[\frac{N}{2}\pi - \left(\frac{N}{2} - 2\right)a\right] + 2g\left[\frac{N}{2}\pi - \left(\frac{N}{2} - 3\right)a\right] \ldots + g\left[\frac{N}{2}\pi + \frac{N}{2}a\right]\right\}. \tag{5}$$

Since $f(\omega t + \theta)$ is periodic in 2π and may be expanded as a Fourier series, $g(\omega t)$ is also such a series, with a change of signs throughout due to the integration. Then if $N/2$ is an even number, Equation (5) becomes:

$$\frac{\Delta W}{\epsilon V} = \frac{1}{a}\left\{2g\left[\frac{N}{2}a\right] - 4g\left[\left(\frac{N}{2} - 1\right)a\right] + 4g\left[\left(\frac{N}{2} - 2\right)a\right]\right.$$
$$\left. \ldots 4g\left[a\right] + 2g\left[0\right]\right\}. \tag{6}$$

(*a*) When $f(\omega t + \theta) = \sin(\omega t + \theta)$, then

$$\frac{\Delta W}{\epsilon V} = \frac{1}{a}\left\{2\cos\frac{N}{2}a - 4\cos\left(\frac{N}{2} - 1\right)a + 4\cos\left(\frac{N}{2} - 2\right)a \ldots - 4\cos a + 2\right\}. \tag{7}$$

This series may most conveniently be summed graphically, and $\Delta W/\epsilon V$ plotted as a function of a. Figs. 2 and 3 show the results for $N = 8$ and $N = 20$ respectively. The abscissa may be converted to a mass scale since $a = km^{\frac{1}{2}}$, where $k = s\omega(2\epsilon V_0)^{-\frac{1}{2}}$.

The mass of the synchronous ion may be determined from the relationship:

$$\frac{M}{n} = 0.0746\frac{V_0}{f^2d^2} \tag{8}$$

where M is the atomic weight of the ion,

n is the multiplicity of the ion charge,

V_0 is the d-c. accelerating voltage (volts),

f is the frequency (Mc. per sec.),

d is the gap spacing (inches).

Some general conclusions are evident:

(1). The maximum value of $\Delta\omega/\epsilon V$ is proportional to the number of radio-frequency gaps, since the height of the main peak is $2N/\pi$.

(2). The height of the main peak for the second cycle of a (i.e., $a \approx 3\pi$) is $2N/3\pi$.

(3). The optimum value of $a \to \pi$ and hence $\theta_{\max} \to 0$ as the number of radio-frequency gaps is increased.

(4). The mass resolution is improved as N increases. This is discussed more fully in Appendix II.

(*b*) When the accelerating wave form is a square wave, then

$$f(\omega t) = \frac{4}{\pi}\left[\sin\omega t + \frac{\sin 3\omega t}{3} + \frac{\sin 5\omega t}{5} + \ldots\right]. \tag{9}$$

Then,

$$(10) \qquad g(\omega t) = -\frac{4}{\pi}\left[\cos \omega t + \frac{\cos 3\omega t}{3^2} + \frac{\sin 5\omega t}{5^2} + \ldots\right]$$

which represents a symmetrical triangular wave form.

All the terms of this series may also be summed graphically and $\Delta W/\epsilon V$ plotted as a function of α for different values of N. Fig. 5 shows the results for $N = 20$. It is evident that the main peak is sharper and higher than for the case of a sinusoidal applied wave form, since

$$\frac{\Delta W}{\epsilon V}(\max) = N.$$

Appendix II

An estimate of the mass resolution of a system with a large number of gaps may be made by finding the energy gained from the radio-frequency field by an ion whose transit angle differs only slightly from the optimum angle α_0. For the general case of N gaps, and for a sinusoidal driving voltage:

$$(11) \quad \Delta W = \frac{vV\epsilon}{s}\left[\sum_{n=1}^{n=m}\int_{2(n-1)s/v}^{(2n-1)s/v} \sin(\omega t + \theta)dt - \sum_{n=1}^{n=m}\int_{(2n-1)s/v}^{2ns/v} \sin(\omega t + \theta)dt\right].$$

Where $n = 1, 2, 3, \ldots\ldots, n$; and $m = \frac{1}{2}N$

$$(12) \qquad W = \frac{V\epsilon}{\alpha}\sum_{n=1}^{n=m}\Big\{\cos[(2n-2)\alpha + \theta] + \cos(2n\alpha + \theta) - 2\cos[(2n-1)\alpha + \theta]\Big\}.$$

Thus,

$$(13) \qquad \frac{\Delta W}{V\epsilon} = \frac{2(\cos\alpha - 1)}{\alpha}\sum \cos[(2n-1)\alpha + \theta].$$

For optimum energy transfer and large N,

$$\alpha_0 \to \pi \text{ and } \theta \to 0.$$

Then,

$$(14) \qquad \frac{\Delta W}{V\epsilon} = \frac{4m}{\pi} = \frac{2N}{\pi}$$

To find the energy transfer at transit angles close to α_0, let $\alpha_1 = \pi + \delta$ where δ is small.

Then from Equation (13)

$$(15) \qquad \frac{\Delta W}{V\epsilon_{(\alpha_1)}} = \frac{2[\cos(\pi + \delta) - 1]}{\pi + \delta}\sum \cos[(2n-1)(\pi + \delta)]$$

$$(16) \qquad = \frac{2(\cos\delta + 1)}{\pi + \delta}\sum \cos\{(2n+1)\delta\}.$$

If $\cos(2n+1)\delta$ is sufficiently small to be replaced by the first two terms of its series expansion, then

$$\sum \cos\{(2n-1)\delta\} = m - \frac{\delta^2}{2}\sum (2n-1)^2$$

$$= m - \frac{\delta^2}{2}\{4\sum n^2 - 4\sum n + m\}$$

$$= \frac{N}{2}\left\{1 - \frac{\delta^2}{6}(N^2-1)\right\}. \tag{17}$$

Combining Equations (16) and (17)

$$\frac{\Delta W}{V\epsilon_{(a_1)}} = \frac{N(\cos\delta + 1)}{\pi + \delta}\left\{1 - \frac{\delta^2}{6}(N^2-1)\right\}$$

for $\delta \to 0$, and N large,

$$\frac{\Delta W}{V\epsilon_{(a_1)}} = \frac{2N}{\pi}\left\{1 - \frac{\delta^2}{6}(N^2-1)\right\}. \tag{18}$$

Then the difference in energy gained at a_1 and a_0 (see Equation (14)) is given by:

$$\frac{\delta W}{V\epsilon} = \frac{N\delta^2}{3\pi}(N^2-1). \tag{19}$$

If the percentage mass resolution is

$$R = \frac{M_1 - M_2}{M_0} \times 100$$

where M_0 is the mass of ion with transit angle a_0, and M_1 and M_2 are the masses of the ions which are just turned back by the retarder grid.

Then,

$$R = \frac{(\pi-\delta)^2 - (\pi+\delta)^2}{\pi^2} \times 100 = -\frac{4\delta}{\pi} \times 100. \tag{20}$$

From Equation (19)

$$\delta^2 = \frac{3k\pi}{N(N^2-1)}$$

where

$$k = \frac{\delta V}{V}$$

and

$$\epsilon\delta V = \delta W.$$

Then

$$R^2 = \frac{48 \times 10^4 k}{N(N^2-1)\pi}$$

or

$$R = 10^2\sqrt{\frac{48K}{N(N^2-1)\pi}}\,\%. \tag{21}$$

R is shown as a function of N for $k = \frac{1}{12}$ in Fig. 4.

PHYSICAL ASPECTS OF THE CONTRACTION HYPOTHESIS OF OROGENESIS[1]

By Adrian E. Scheidegger

ABSTRACT

Several geophysical arguments important in the theoretical consideration of the contraction hypothesis of orogenesis are analyzed in detail with respect to their physical bases. Specifically, it is shown (*a*) that a thin surface layer of a sphere will fold in a smaller pattern than a thicker surface layer if the interior shrinks by the same amount; (*b*) that the speed of cooling of a sphere decreases if some radioactive and hence heat generating material is moved from the surface towards the center (and conversely); and (*c*) that at the level of no strain of a cooling sphere the temperature–depth curve must be almost straight. Those geological theories in which these considerations are important are indicated.

1. INTRODUCTION

It is noted that the earth is a heat engine. Several theories have been developed for finding a relation between the thermodynamic facts that are known about the earth and its surface features, such as oceans, continents, and mountain belts.

According to many theories the earth was at one stage of its life a hot, liquid body, which quickly solidified and continued to cool by radiation. Discrepancies arise with respect to the causes of orogenesis, that is mountain building. Some claim that slow thermal convection currents in a highly viscous mantle are the chief cause of orogenesis, but others assume that contraction due to cooling of the interior of the earth has that same effect.

To discuss these theories from both geological and physical viewpoints a geologist, J. T. Wilson, and a physicist, the writer, have studied the problem together and written a joint paper (6). The present paper is a theoretical discussion of some of the basic physical principles that are involved in the theory of the history of the earth that was developed there.

The National Research Council of Canada provided funds to enable the writer to pursue most of these investigations during the summer of 1950 in Toronto. They were completed at Queen's University.

Thus it should be understood that the present author is not responsible for the geological ideas contained in this paper. The object is merely to discuss from a physical point of view some of the theories that have been advanced in explanation of some characteristic features of the earth. It is a collection of theoretical discussions of specific questions, which occur in connection with a theory of the thermal history of the earth, rather than an account of such a theory. It is believed, however, that this paper, together with others written recently at Toronto (1, 6–11), all of which are based upon the fundamental contraction theory as advanced by Jeffreys (5), form a scientific structure of

[1] *Manuscript received May 17, 1951.*
Contribution of the Department of Physics, Queen's University, Kingston, Ont., and of the Geophysics Laboratory, University of Toronto, Toronto, Ont.

considerable strength which may be capable of explaining the principal causes of the earth's orogenetic movements. The reasons for advocating the contraction hypothesis have been stated in a previous paper (6).

The contraction theory as stated by Jeffreys (5) is that the mantle solidified outwards from its base at the top of the liquid iron core and has since been cooling by conduction without convection currents. From the center of the earth to within about 700 km. of the surface there has not been time since the earth solidified for any appreciable cooling or change in volume to have taken place. Within the region from about 700 to 70 km., cooling by conduction is taking place and hence this shell or layer is contracting and being stretched about the unchanging interior. Hence it is in a state of internal tension.

Near the surface the rocks have already largely cooled so that at the surface they are in thermal equilibrium with the heat provided by solar radiation. They are therefore not changing very much in temperature and the cooling and contraction of the layer or shell beneath them puts the outermost shell into a state of internal compression above a level of no strain at 70 km. depth (see Reference 11).

To summarize we may thus note that the contraction hypothesis upon grounds of thermal and mechanical behavior divides the earth into three shells. These shells are not dependent upon the earth's composition and hence should not be confused with such terms as core, mantle, and crust. As the earth cools the boundaries between the shells move deeper into the earth.

The problem for which Wilson and the writer previously (6) attempted a solution was to find what physical explanation might be advanced to explain the failure of the outermost shell.

It has long been known that the active zones of the earth lie along the Alpine–Himalayan Mountain chains and in a belt around the Pacific. These belts are where failure is occurring. It has also been recognized that many island and mountain arcs which make up these belts are approximately circular in shape. The problem was therefore to discover some mechanism which could explain the formation of individual arcs and which would allow a number of arcs to be formed together so as to constitute belts. The solution proposed was that contraction below the level of no strain is uniform in all directions and that this causes failure by sliding fracture. It was also supposed that there are weak points upon the spherical shells and that these failures are shaped like parts of cones and give arcuate intersections with the spherical surfaces having these weak points at their centers.

Examination of the seismological and geological evidence has shown that earthquakes which give evidence of shear failure do occur arranged upon such conical surfaces or zones to depths of 700 km. (3) and that the centers of many arcs lie at the margins of continents or near these former margins where they might be assumed to have been weakened by deposition of deltaic loads of detritus. The discovery that earthquakes do occur to depths as great as 700 km. is either a confirmation of Jeffreys' prediction of cooling to that depth or else a remarkable coincidence.

Along these belts of failure volcanoes occur. These and other evidence of igneous activity show that the fractures are channels by which material escapes from within the earth. The evidence suggests that the effusive material tends to be lighter and more siliceous than the average rock occurring deep within the earth. Geological evidence supplemented by radioactive age determinations also suggests that the pattern of failure has moved from time to time and that successive failures each contributing siliceous rock material have caused the growth of continents outwards from nuclei and that this is still continuing.

It will be our task to discuss in the following sections the soundness of some of the physical principles on which the geological theory outlined above is founded. In doing so, we shall take up three points one after another and investigate what can be said about each from the standpoint of a physicist.

2. THE THICKNESS OF THE EARTH'S CRUST AND MOUNTAIN BUILDING

The first problem discussed here concerns the relative size of present and of very ancient mountains. Presumably the depth to which cooling has penetrated and also the level of no strain have both been moving inwards during geological time. Therefore, the crust involved in folding and, hence, it will be shown, the size of the folds were smaller during the early history of the earth and have grown larger in succeeding periods. There is some geological evidence that this is so. In particular, the pattern of folds in Keewatin or earliest Pre-Cambrian time as shown on geological maps of Canada shows a quite different pattern from mountain ranges of younger age.

It is a natural suggestion that the different pattern of the Pre-Cambrian mountain belts is due to the deepening of the level of no strain, which is supposed to have taken place in the meantime.

We shall investigate now whether such an assumption can be confirmed by more theoretical considerations of the types of folding that must occur when the outermost shell above the level of no strain is supposed to become thicker.

Thus, the model of the mechanism of mountain building which we are to consider is as follows.

Below a surface shell of constant density we assume a contracting spherical interior. If the contraction takes place at a constant rate, the question arises as to what extent folding depends on the thickness of the surface shell.

Let the volume of the shell be V and the inner radius of this shell be a and its thickness b. If the shell is a relatively thin and cool layer, it will retain its volume almost unchanged during any short period of contraction brought about by cooling in the larger and hotter interior. Then we have

$$V = 4\pi a^2 b. \tag{2.1}$$

If the radius of the interior sphere is changed by the amount da and the thickness of the shell by db, then we obtain

$$dV = 4\pi a^2 db + 8\pi ab\, da. \tag{2.2}$$

However, dV must vanish, as stated above, hence

$$2da\,b + a\,db = 0, \tag{2.3}$$

$$db = -2b\,da/a. \tag{2.4}$$

This equation indicates that, for a shrinking of the interior by the amount da, the amount of material of the shell that has to be moved in order to make it continue to fit the interior is proportional to its thickness.

If we assume that the shrinking of the interior occurs at a constant rate $\dot{a}$ in time, then (2.4) becomes

$$db/dt = -(2b/a)\dot{a}, \tag{2.5}$$

which indicates that the material in the shell that has to be moved around per unit time is just proportional to the thickness of the shell.

This shows indeed that a thin shell must form smaller mountain ranges than a thick one, if the interior shrinks by the same amount. Thus, if the Keewatin folds were formed in that manner, they would be expected to be closer together than later folds formed in a similar way.

We may corroborate this statement and estimate how much the mantle has to cool so as to cause a major world-wide orogenetic system. It is impossible to give any accurate estimate of the quantities involved in mountain building, as most of them are so imperfectly known that one can only estimate the order of magnitude.

Thus, in an orogenetic diastrophism one may assume that two-thirds of two great circles about the earth are folded into mountains 2000 m. high and 300 km. wide. The volume increase is thus $(2/3)\times 2\times 40{,}000\times 2\times 300$ km.^{3} or 32×10^6 km.^{3} This corresponds to an over-all increase in thickness of the outermost shell of $(32\times 10^6)/(45\times 10^7) = 7\times 10^{-2}$ km. It follows from Equation (2.4) that in order to obtain an increase of thickness of $db = 7\times 10^{-2}$ km., assuming $b = 70$ km., one has

$$da/a = -\tfrac{1}{2}\,db/b = -\tfrac{1}{2}\times\frac{7\times 10^{-2}}{7\times 10} = -5\times 10^{-4}. \tag{2.6}$$

Thus one needs an over-all relative linear contraction of 0.05% in the interior to fold up these mountains. The radius of the interior of the earth is thereby contracting by about 3 km.

We shall assume now that the cooling and contraction takes place in depths between 70 and 700 km. only. Thus the thickness of that layer is changed from 630 to 627 km. during one orogenetic cycle. We can calculate the relative volume contraction which has therefore to occur; it is

$$\frac{8\pi\times(5930)^2\times 3}{\frac{4\pi}{3}\{(5930)^3-(5300)^3\}} = .0106. \tag{2.7}$$

The relative linear contraction is one-third of that, and thus is equal to 0.0035 or 0.35%. From the Handbook of Physical Constants (4) we see that the relative linear expansion of granite equals about 10^{-5} per degree Centigrade, so that an

over-all change in temperature of 350° C. in the cooling layer will cause sufficient contraction to give rise to the mountains.

This seems quite reasonable. If the temperature of the liquid iron core is assumed to be about 3500° C., then 10 orogenetic cycles would be possible while a particular part of the mantle cooled down from the temperature of the core to the temperature of outer space.

3. RADIOACTIVITY AND RATE OF COOLING OF EARTH

As mentioned in the introduction to this paper, a theory of the growth of continents has been recently postulated. It includes the argument that at the edges of existing continents breakages occur in the outer shell of the earth which give rise to island arcs. The space between the continent and the island arc is filled up by sedimentary material and by new volcanic rocks and thus gradually becomes incorporated into the continent. Then this procedure starts anew. An important part of this theory is that the radioactive elements rise when fracturing occurs and thus become concentrated in the continental material. This should cause a hastening of the cooling of the earth and also have some mechanical influence on the breakages. If this factor was not important the time between orogenetic periods might be expected to have increased with geological time, and this does not appear to have been the case.

We will therefore investigate whether this assumption is in agreement with the laws of heat conduction. It seems reasonable that the heat escaping from the earth comes mainly from the outer layers and that the deeper parts do not contribute appreciably to that. Thus, there is little doubt that if radioactive material is moved upward the speed of cooling of the earth increases, and conversely. However, no actual proof of that is known to us. Because of the importance of the assumption in question in Wilson's papers (7-11) it seems desirable to attempt the proof here.

(*a*) *General Theory*

In order to investigate the change of temperature θ in a body we have to solve the equation

$$-\dot{\theta} + \beta\Delta\theta + \rho = 0. \tag{3.1}$$

Herein, β is a constant coefficient and ρ the rate of creation of heat; i.e. ρ is proportional to the radioactivity. It shall be assumed in our model that ρ is a function of the space coordinates only. If we assume that the initial temperature is θ_0 and that there is radiation at the surface we have to observe the following boundary conditions.

$$\theta(X, Y, Z, t = 0) = \theta_0, \tag{3.2}$$

$$\text{grad}_n\, \theta \text{ (surface)} = -\alpha\theta, \tag{3.3}$$

where α is a positive constant of the material. Equation (3.3) is Newton's law of cooling. One has to solve first the homogeneous equation of (3.1), namely

$$-\dot{\theta} + \beta\Delta\theta = 0 \tag{3.4}$$

under the same boundary conditions as above (3.2), (3.3). The solution is obtained in the usual manner by assuming

$$\theta = T(t)\, R(X, Y, Z). \tag{3.5}$$

This leads to

$$\dot{T}/T = \beta\Delta R/R = -\lambda, \tag{3.6}$$

where the eigenvalues λ have to be determined so that the boundary condition (3.3) is satisfied. The latter leads to the eigenvalue equation. Then the general solution of (3.4) is

$$\theta = \sum_{\lambda} c_\lambda T_\lambda R_\lambda. \tag{3.7}$$

The coefficients c_λ are determined so that (3.2) is fulfilled.

With the eigenfunctions R_λ one can attempt the solution of the inhomogeneous equation (3.1) in the following way

$$\theta = \sum_{\lambda} \gamma_\lambda(t)\, R_\lambda(X, Y, Z). \tag{3.8}$$

This gives inserted into the differential equation (3.1)

$$\sum \dot{\gamma}_\lambda R_\lambda + \beta \sum \gamma_\lambda \Delta R_\lambda + \rho = 0, \tag{3.9}$$

or, by means of (3.6)

$$-\sum \dot{\gamma}_\lambda R_\lambda - \sum \gamma_\lambda R_\lambda \lambda + \rho = 0. \tag{3.10}$$

Multiplied by R_μ and integrated over the body concerned (because of the orthogonality properties of the eigenfunctions)

$$-A_\mu \dot{\gamma}_\mu - A_\mu \gamma_\mu \mu + \int \rho R_\mu \, dV = 0, \tag{3.11}$$

where A_μ are the normalization constants

$$A_\mu = \int R_\mu^2 \, dV. \tag{3.12}$$

Equation (3.11) can be written as follows:

$$\dot{\gamma}_\mu + \gamma_\mu \mu = \int \rho (R_\mu / A_\mu) \, dV, \tag{3.13}$$

whereof the general solution is

$$\gamma_\mu(t) = B_\mu \exp(-\mu t) + (1/\mu) \int \rho (R_\mu / A_\mu) \, dV. \tag{3.14}$$

B_μ is determined by the initial conditions. Hence

$$\theta = \sum_{\lambda} \{B_\lambda \exp(-\lambda t) + (1/\lambda) \int \rho (R_\lambda / A_\lambda) \, dV\} R_\lambda(X, Y, Z). \tag{3.15}$$

We are interested in the cooling, i.e. in the quantity

$$\frac{d}{dt} \int \theta \, dV = \int \dot{\theta} \, dV. \tag{3.16}$$

For given initial conditions we have a certain set of functions

$$\gamma_\lambda(0) = \gamma_{\lambda 0}. \tag{3.17}$$

Hence

$$B_\lambda = \gamma_{\lambda 0} - \frac{1}{\lambda}\int \rho(R_\lambda/A_\lambda)\, dV. \tag{3.18}$$

Inserted into (3.16)

$$\int \theta\, dV = \sum_\lambda \{-\gamma_{\lambda 0}\lambda \exp(-\lambda t)\int R_\lambda\, dV + \exp(-\lambda t)\int \rho \frac{R_\lambda}{A_\lambda}\, dV \int R_\lambda\, dV\}. \tag{3.19}$$

In (3.19) only the second term depends on the radioactivity. The rate of cooling increases if the second term (which we denote by H) becomes smaller.

(*b*) *Example of a Sphere*

The general method above is now to be applied to the case of a sphere. A straightforward calculation yields the following eigenfunctions,

$$R_\lambda = \frac{1}{r}\sin \sqrt{\frac{\lambda}{\beta}}\, r.$$

The eigenvalues are subject to the eigenvalue equation

$$-\frac{1}{a} + \sqrt{\frac{\lambda}{\beta}}\cot \sqrt{\frac{\lambda}{\beta}}\, a = -a. \tag{3.20}$$

In order to find the rate of cooling in dependence from the distribution of the radioactivity we assume that the total radioactivity P is concentrated in one shell of radius b; thus

$$\rho = \frac{P}{4\pi r^2}\delta(r - b), \tag{3.21}$$

(δ being the Dirac function).

Thus the dependence $-H$ of the rate of cooling of a sphere on such a radioactive shell is

$$H = \sum_\lambda \rho \frac{R_\lambda}{A_\lambda}\, dV \int R_\lambda dV = \sum_\lambda \frac{1}{A_\lambda} P \frac{1}{b}\sin\left(\sqrt{\frac{\lambda}{\beta}}\, b\right) 4\pi a \frac{\beta}{\lambda} a \sin\left(\sqrt{\frac{\lambda}{\beta}}\, a\right) e^{-\lambda t}. \tag{3.22}$$

We are interested in whether this rate decreases when some of the radioactive material is moved from the surface towards the center. The rate of cooling decreases if

$$\frac{\partial H}{\partial b} < 0. \tag{3.23}$$

The differentiation is legitimate since the resultant series converges for any value $t > 0$. If the radioactive material is taken from the surface ($b = a$) we obtain

$$(3.24)\quad \frac{\partial H}{\partial b} = \sum_{\lambda} \frac{P}{A_\lambda}\left[-\frac{1}{a^2}\sin\left(\sqrt{\frac{\lambda}{\beta}}a\right) + \frac{1}{a}\sqrt{\frac{\lambda}{\beta}}\cos\left(\sqrt{\frac{\lambda}{\beta}}a\right)\right] \times 4\pi a \frac{\beta}{\lambda} a \sin\left(\sqrt{\frac{\lambda}{\beta}}a\right) e^{-\lambda t}.$$

The square bracket in (3.24) can be rewritten using the eigenvalue equation (3.20). It follows from (3.20)

$$\frac{1}{a}\sqrt{\frac{\lambda}{\beta}}\cos\sqrt{\frac{\lambda}{\beta}}a = \frac{1}{a^2}\sin\sqrt{\frac{\lambda}{\beta}}a - \frac{a}{a}\sin\sqrt{\frac{\lambda}{\beta}}a.$$

Hence

$$(3.25)\quad \frac{\partial H}{\partial b} = \sum_{\lambda} -\frac{4\pi}{A_\lambda} P \frac{\beta}{\lambda} a^2 \sin^2\left(\sqrt{\frac{\lambda}{\beta}}a\right) e^{-\lambda t} < 0,$$

since A_λ is essentially positive.

This is the general proof of the assumption that the speed of cooling of a sphere decreases if some radioactive material is moved from the surface towards the center (and conversely).

It would be interesting to know how large the effect of the change of position of radioactivity is. One might think that one could calculate this directly from Equation (3.25). This is, however, not the case. The physical constants which occur in that equation are not well known so that the result turns out to be quite arbitrary.

4. CONNECTION BETWEEN STRESSES AND TEMPERATURE IN THE EARTH

An important problem in the contraction theory of orogenesis is to justify the assumed distribution of stresses in the earth. Since the stresses are believed to be due to thermal reasons, there must be a connection between the temperature curve and the stresses within the earth. Of especial importance is the question of whether the assumption of a level of no strain in the neighborhood of 70 or 100 km. depth is compatible with the data known about the temperature distribution. The contraction hypothesis stands or falls upon the possibility of the existence of such a level of no strain at or near the specified depth.

We note that already Jeffreys (5, p. 280), had stated a connection between the stresses and the temperature in the earth. His procedure was as follows.

Consider a shell of internal radius r and thickness δr. Let its coefficient of linear expansion be n, and the density ρ. Let the rise of temperature of the shell be dT. Then dT is a function of r alone. The density of the shell, if we ignore the small change due to compressibility, becomes $\rho(1 - 3n\,dT)$. Let the radius of the shell of radius r become $r(1 + da)$. Then the external radius becomes

$$(4.1)\quad r(1 + da) + \delta r\left\{1 + \frac{d}{dr}(r\,da)\right\}.$$

Hence the mass of the shell after the change of temperature is

$$4\pi r^2(1+da)^2\delta r\left\{1+\frac{d}{dr}(r\,da)\right\}\rho\{1-3n\,dT\}$$
$$=4\pi\rho r^2\delta r\left\{1+2\,da+\frac{d}{dr}(r\,da)-3n\,dT\right\}, \tag{4.2}$$

neglecting the squares and products of da and dT. But the mass is unaltered. Hence we have the equation of continuity:

$$2da+\frac{d}{dr}(r\,da)-3n\,dT=0. \tag{4.3}$$

Since dT is supposed to be known throughout the earth, this is a differential equation to determine da.

Now if a shell simply expanded without stretching, its radius would increase by $rn\,dT$ instead of $r\,da$ so that the stretching required to make it continue to fit its interior is $r(da-n\,dT)=r\,dS$. The expression dS is thus immediately connected with the strains. Then this gives us

$$(d/dr)(dS\,r^3)=-r^3(d/dr)(n\,dT), \tag{4.4}$$

$$dS=-\frac{1}{r^3}\int_0^r r^3\frac{d}{dr}(n\,dT)\,dr, \tag{4.5}$$

or

$$dS=-n\,dT+\frac{1}{r^3}\int_0^r 3r^2n\,dT\,dr. \tag{4.6}$$

The differential dT has to be understood with respect to time. Thus we find

$$\frac{\partial S}{\partial t}=-n\frac{\partial T}{\partial t}+\frac{1}{r^3}\int_0^r 3r^2n\frac{\partial T}{\partial t}dr. \tag{4.7}$$

So far we have been following Jeffreys (5). But now we note that we can substitute $\partial T/\partial t$ from the heat conduction equation

$$\rho c\frac{\partial T}{\partial t}=k\Delta T+p. \tag{4.8}$$

In spherical coordinates this is (if T depends on r only)

$$\rho c\frac{\partial T}{\partial t}=k\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial T}{\partial r}\right)+p. \tag{4.9}$$

Hence

$$\frac{\partial S}{\partial t}=-\frac{n}{\rho c}\left[k\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial T}{\partial r}\right)+p\right]+\frac{3}{r^3}\int_0^r\frac{nr^2}{\rho c}\left[k\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial T}{\partial r}\right)+p\right]dr, \tag{4.10}$$

or:

$$\frac{\partial S}{\partial t}=-\frac{n}{\rho c}\left[\frac{k}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial T}{\partial r}\right)+p\right]+\frac{3}{r^3}\int_0^r\frac{nk}{\rho c}\frac{\partial}{\partial r}\left(r^2\frac{\partial T}{\partial r}\right)dr+\int_0^r\frac{n}{\rho c}pr^2\,dr. \tag{4.11}$$

Now, the object is to evaluate the integrals in Equation (4.11) from the center of the earth up to a point r somewhere in the mantle. Thus, we have to integrate over the whole core, across the discontinuity between core and mantle,

and finally over part of the mantle. Let us denote the radius of the core by R_0 (= 2900 km.), then we have a discontinuity in the integration for $r = R_0$. Before (undashed) and after (dashed) this level, n, k, ρ and c are assumed to be constant. Furthermore, we neglect the radioactivity p. Hence, for $r > R_0$:

$$\frac{\partial S}{\partial t} = -\frac{n'}{\rho' c'}\left[k'\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial T}{\partial r}\right)\right] + \frac{3}{r^3}\left\{\frac{nk}{\rho c}\left(r^2\frac{\partial T}{\partial r}\right)\Big|_0^{R_0} + \frac{n'k'}{\rho' c'}\left(r^2\frac{\partial T}{\partial r}\right)\Big|_{R_0}^{r}\right\}. \tag{4.12}$$

Thus

$$\frac{\partial S}{\partial t} = -\frac{n'}{\rho' c'}\left[k'\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial T}{\partial r}\right)\right] + \frac{3}{r^3}\left\{\left(\frac{nk}{\rho c} - \frac{n'k'}{\rho' c'}\right)R_0^2\frac{\partial T}{\partial r}\Big|_{R_0} + \frac{n'k'}{\rho' c'}r^2\frac{\partial T}{\partial r}\right\}. \tag{4.13}$$

The term from the discontinuity vanishes if we assume that the temperature gradient for R_0 vanishes. This may be assumed, as we can believe that the temperature is roughly constant at that depth of 2900 km. Then we have for r larger than R_0:

$$\frac{\partial S}{\partial t} = \frac{n'}{\rho' c'}k'\frac{1}{r}\frac{\partial T}{\partial r} - \frac{n'}{\rho' c'}k'\frac{\partial^2 T}{\partial r^2}. \tag{4.14}$$

This is the equation connecting stress and temperature in the earth. In order to evaluate this equation a little more, we may ask what is the condition for the temperature curve so that a level of no strain can exist in the mantle. The condition for a level of no strain is according to Jeffreys (5) $\partial S/\partial t = 0$, hence

$$\frac{1}{r}\frac{\partial T}{\partial r} = \frac{\partial^2 T}{\partial r^2}. \tag{4.15}$$

If we have any given temperature distribution within the earth, then the layers in which the above equation is satisfied are levels of no strain. It is obvious that the position of the level of no strain has nothing to do with the coefficients of heat conduction, etc., but is determined wholly by the *shape* of the temperature curve with depth.

In spite of these remarks, Jeffreys claims to have determined the position of the level of no strain from thermodynamic reasons. In his calculations he does not replace $\partial T/\partial t$ by the space derivatives. In our opinion, this makes the whole situation much worse, because now one has to know the present *rate* of cooling for each layer of the earth. If the coefficient of expansion is taken to be constant (at least in the mantle), then it drops out from the equation determining the level of no strain. Hence it is entirely incomprehensible to us how it should be possible to connect the level of no strain with the thermal constants of the earth, and how Jeffreys arrived at his 100 km. from them.

Jeffreys approximated Equation (4.7) by the assumption of a plane surface of the earth. We can do the same thing here. Putting x = depth, i.e.

(4.16) $$r = R - x,$$

and making the same assumptions as Jeffreys, Equation (4.7) becomes

$$\frac{\partial S}{\partial t} = -n\frac{\partial T}{\partial t} + \frac{3}{R}\int_x^\infty n\frac{\partial T}{\partial t}dx.$$

In this case, we have

(4.17) $$\rho c\frac{\partial T}{\partial t} = k\frac{\partial^2 T}{\partial x^2};$$

hence

(4.18) $$\frac{\partial S}{\partial t} = -\frac{nk}{\rho c}\frac{\partial^2 T}{\partial x^2} + \frac{3}{R}\int_x^\infty \frac{nk}{\rho c}\frac{\partial^2 T}{\partial x^2}dx.$$

For the level of no strain this must vanish, hence

(4.19) $$\frac{\partial^2 T}{\partial x^2} = -\frac{3}{R}\frac{\partial T}{\partial x}.$$

This is the condition that corresponds to our Equation (4.15) in Jeffreys' approximation. It is not understood why Jeffreys did not consider the exact solution, since it is not more difficult to obtain it than the approximation (4.16). The two solutions are obviously somewhat different so that it does not seem quite legitimate to assume that the earth is flat.

We can investigate how well several assumptions of temperature curves fit the concept of a level of no strain at about 70 km. depth as usually postulated in the contraction hypothesis. The gradient is at those depths about 1°–1½°/km. (see Gutenberg (2)). Then, Formula (4.15) gives us

(4.20) $$\partial^2 T/\partial r^2 = (1/6000) \times 1\tfrac{1}{2} = .00025\ ^\circ/\text{km}^2.$$

The change of the gradient over a distance of 100 km. in the level of no strain becomes

(4.21) $$(\partial^2 T/\partial r^2)\delta r = .00025 \times 100 = .025\ ^\circ/\text{km}.$$

This means that, over a distance of 100 km., the gradient should change in the level of no strain by about 1/40°/km., which is almost negligible. This indicates that in the level of no strain the temperature curve must be almost straight. This would still be the case if the temperature gradient were assumed to be much larger.

Thus, we are able to see how well several proposed temperature curves fit the concept of a level of no strain at a 70 km. depth. Jeffreys (5, Fig. 13) assumes a line with fairly strong curvature in the outer parts of the earth (down to about 100 km.) and then a rather flat curve with again an increase in curvature at a greater depth. In other words, he assumes as a temperature curve a line that consists of three parts, a curved section, a flat section, and again a curved section. The level of no strain is where the curved section joins the flat one.

The temperature curves of several other authors are chosen in such a way that the flat part sets in at a depth of about 70 km. where one would like to have the level of no strain because of the contraction hypothesis. (cf. Gutenberg(2, p. 162)). The suggestion of Gutenberg himself, however, to assume a strong deflection in the temperature curve at about 70 km. seems substantially different as this must cause great stresses. At any rate, it would not suit the contraction hypothesis very well.

ACKNOWLEDGMENTS

In conclusion, I should like to acknowledge my indebtedness to Dr. J. Tuzo Wilson who suggested the scope of this paper. Without his unceasing interest and many valuable suggestions these investigations would never have been completed. Thanks are also due to Mr. J. R. Pounder of the University of Toronto who checked the calculation of Section 3, and to the National Research Council of Canada which sponsored the present work.

REFERENCES

1. BULLARD, E. C. Monthly Notices Roy. Astron. Soc. Geophys. Suppl. 6: 36. 1950.
2. GUTENBERG, B. Internal constitution of the earth. McGraw-Hill Book Co., Inc., New York. 1939.
3. GUTENBERG, B. and RICHTER, C. F. Seismicity of the earth. Princeton University Press, Princeton, N.J. 1949.
4. Handbook of Physical Constants. *Edited by* F. Birch. Geological Society of America special papers No. 36. 1942.
5. JEFFREYS, H. The earth. Cambridge Univ. Press, London. 1929.
6. SCHEIDEGGER, A. E. and WILSON, J. T. Proc. Geol. Assoc. Canada, 3: 167. 1950.
7. WILSON, J. T. Trans. Am. Geophys. Union, 29: 691. 1948.
8. WILSON, J. T. Can. Mining Met. Bull. 42: 543. 1949.
9. WILSON, J. T. Nature, 164: 147. 1949.
10. WILSON, J. T. Trans. Roy. Soc. Can. IV, 157. 1949.
11. WILSON, J. T. Proc. Geol. Assoc. Can. 3: 141. 1950.

DIELECTRIC BEHAVIOR AT THE SECOND-ORDER TRANSITIONS IN CHROMIUM SULPHATE, CHROMIUM NITRATE, AND NICKEL NITRATE[1]

By H. Grayson-Smith[2] and R. F. Sturrock[3]

ABSTRACT

In order to obtain further evidence concerning the nature of the second-order transitions which were discovered by Vasileff and Grayson-Smith in certain salts of the iron group metals, dielectric measurements have been made at temperatures down to about 100° K. In the case of $Cr_2(SO_4)_3.18H_2O$ the transition at 195° K. has a marked effect on the dielectric properties, which is clearly due to the freeing of polar molecular groups, so that they can be oriented by an electric field. The entropy of transition is close to $R \ln 4$, which suggests that two groups per molecule are set free, to take either of two quantized orientations. The dipole moment of the group which is set free can be estimated from the temperature dependence of the dielectric constant just above the transition point and has been found to be 2.17×10^{-18} e.s.u., approximately equal to the dipole moment of the H_2O molecule. It is likely, therefore, that the second-order transition in this salt is caused by the freeing of two of the 18 H_2O groups. For chromium and nickel nitrates the transitions have no observable effect on the dielectric properties, and it is suggested that in these salts NO_3 groups are set free.

INTRODUCTION

Specific heat measurements made by Vasileff and Grayson-Smith (6) have shown that several salts of the iron group metals have low temperature second-order transitions, as listed in Table I. The cause of these transitions could not be

TABLE I

SECOND-ORDER TRANSITIONS IN IRON GROUP SALTS

Salt	Transition temperature, ° K.
$Cr_2(SO_4)_3.18\ H_2O$	195°
$Cr(NO_3)_3.9H_2O$	158°
$Co(NO_3)_2.6H_2O$	153°
$Ni(NO_3)_2.6H_2O$	149°

determined from the calorimetric evidence alone. It was known that, although the materials are all paramagnetic, the magnetic properties are perfectly regular down to 65° K. (2). It was suspected, therefore, that the transitions were due to the onset of rotation of molecular groups within the crystal. The purpose of the present investigation was to see whether further evidence concerning the transitions could be obtained from a study of the dielectric properties. Dielectric measurements have therefore been made, down to about 100° K., with hydrated chromium sulphate, chromium nitrate, and nickel nitrate. Anhydrous chromium sulphate, which has no transition in this temperature range, was also measured for comparison.

[1] *Manuscript received June 25, 1951.*
Contribution from the Physics Laboratory of the University of Alberta, Edmonton, Alta.
[2] *Professor of Physics, University of Alberta, Edmonton.*
[3] *Now at the Physical Laboratory of the Bureau of Mines, Ottawa.*

MATERIALS USED

The salts were obtained from the same source (British Drug Houses (Canada) Ltd.) as those used by Vasileff and Grayson-Smith, and had the same purity specifications, but were not necessarily from the same batches. The anhydrous chromium sulphate was prepared from the hydrate by heating at about 400° C. until the weight was constant. For three samples the average loss of weight in dehydration amounted to 16.7 H_2O per molecule of $Cr_2(SO_4)_3$. It is evident, therefore, that some dehydration of the original sample had taken place.

When the nitrates were exposed to the air for a few minutes, while the larger grains were being broken, visible traces of moisture accumulated. These salts had to be kept in a desiccator over sulphuric acid for two or three days before they could be used. The dry powder was then transferred as quickly as possible to the experimental cell and sealed to protect it from the air. The chromium sulphate, either hydrated or anhydrous, did not absorb moisture nearly so rapidly and could be transferred to the cell immediately after grinding.

METHOD OF MEASUREMENT

Since the information required was mainly the temperature dependence of the dielectric constant, and there seemed to be no great interest in obtaining accurate absolute values, it was decided that measurements of the powdered salts would be sufficient. For this purpose the salts were packed into the space between two concentric brass tubes, forming a cylindrical condenser of the following dimensions:

Outer diameter - - - 4.242 cm. Inner diameter - - - 3.812 cm.
Height - - - - - - 8.781 cm. Volume - - - - - - - 23.88 cc.
Capacitance empty - - 45.65 $\mu\mu$f.

To hold the cylinders in position they were threaded into lucite rings, with a little liquid solder on the threads to seal the cell and protect the salts from moisture.

The condenser cell was suspended over boiling liquid air in a tall vacuum flask, by means of a long brass tube which passed through a hole in the cover of the flask, and could be fixed at any level by means of a set screw. A polished copper can, closed at the bottom, was placed around the cell, to reduce the vertical temperature gradient and to prevent any possibility of liquid air seeping into the space between the cylinders. By lowering the cell assembly into the vacuum flask a little at a time, until finally the lower portion of the protecting can was immersed in the liquid air, the temperature could be reduced very slowly from room temperature to about 100° K. By raising the cell towards the top of the flask, and then allowing the liquid air to boil away, the temperature could be slowly increased again. Normally, the cooling process extended over about eight hours, and the return to room temperature about six hours, so that many readings could be taken at closely spaced temperature intervals, and the temperature could safely be treated as constant during the few seconds required to make a capacitance reading.

The temperature of the condenser cell was measured by means of two copper–constantan thermocouples held by spring clips against the inner and outer walls. The thermal e.m.f.'s were measured with a potentiometer, and the thermocouples were calibrated at the sublimation point of carbon dioxide and at the boiling point of oxygen. As might be expected with the simple method of temperature control, there was a difference of temperature between the outer and inner thermocouples, amounting at times to as much as three degrees. It was found that the temperature of the inner thermocouple changed very regularly as the cooling or warming proceeded, while that of the outer thermocouple was sometimes erratic. The temperatures used as abscissae in the graphs are those of the inner thermocouple, and may be as much as one degree lower than the mean temperatures of the salt samples. This is unimportant for the purpose of the experiment, since the transition temperatures of the salts are known from the work of Vasileff and Grayson-Smith.

The capacitance C_x of the cylindrical condenser containing the salt was measured by means of a specially constructed Wien bridge, which is shown in Fig. 1. The supporting tube served as the lead to the outer cylinder of the

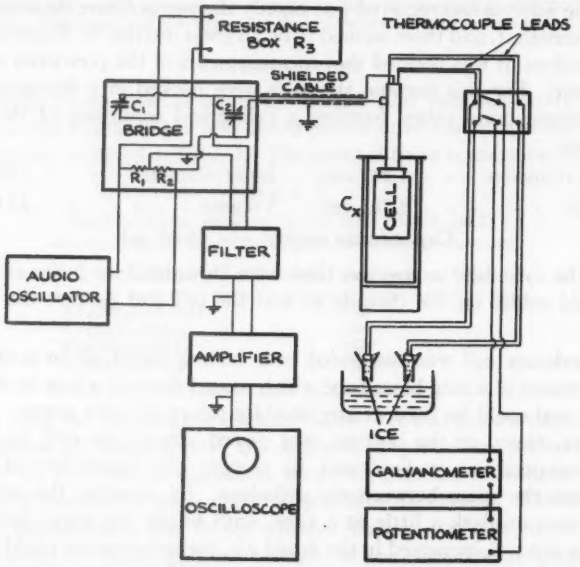

FIG. 1. Circuit for measurement of dielectric constant at low temperatures.

condenser, while the lead to the inner cylinder passed through the tube, forming a concentric cable. In this circuit C_x was in parallel with the calibrated variable condenser C_2, and changes of C_x with temperature were always measured by adjusting C_2 to restore the previous balance. Effectively, therefore, a substitution method was used. Towards the upper end of the temperature range

covered in the measurements, the salts all caused a large imaginary component in the reactance, indicating a large dielectric absorption. This was compensated by adjusting the resistance box R_3. When dielectric absorption occurred it was not possible to balance all harmonics of the oscillator frequency simultaneously. To avoid this difficulty a low pass filter was inserted between the bridge and the oscilloscope which was used as a detector. The filter was designed to pass only the fundamental of the oscillator frequency, 10 kilocycles per second, and all measurements were made at this frequency.

REDUCTION OF THE MEASUREMENTS

When a powdered salt is used, the dielectric constant obtained from the capacitance measurements is that of a mixture of salt and air. In order to derive absolute values of the dielectric constant of the salt it is necessary to know the density accurately, and to make some assumption concerning the shape of the grains. The densities are difficult to determine, and for all the salts in question there are large discrepancies between the values reported by different observers. Therefore, no attempt was made to derive the dielectric constants of the salts. Instead, the results have been expressed in terms of an effective molecular polarizability, derived directly from the measured dielectric constants of the powders.

There is still a great deal of uncertainty concerning the correct expression for the dielectric constant, K, of a dense mixture in terms of the polarizabilities, a_i, of the constituent molecules. The classic expression is the Lorenz-Lorentz formula,

$$\frac{K-1}{K+2} = \frac{4\pi}{3}\sum n_i a_i, \tag{1}$$

where n_i is the number per cc. of molecules of the kind i. In the mixture of salt and air, the number n_1 of salt molecules is much greater than any other n_i, and the effect of the air can be neglected. Moreover $n_1 = mN_0/Mv$, where m is the mass of salt used, M the molecular weight, N_0 Avogadro's number, and v the volume of the cell. Thus the density of the salt does not enter explicitly.

However, Formula (1) does not strictly apply to solids, especially if there are polar molecules present which can change their orientation and, as will appear below, the results of the present experiment show that this is probably the case for hydrated chromium sulphate. Fröhlich (1) has discussed the various methods of approach to the problem. Of these, the one most appropriate to chromium sulphate is due to Onsager (4), who took into account explicitly the dipole–dipole interaction in a condensed isotropic medium, and obtained the result,

$$\frac{(K-1)V}{4\pi} = \frac{3KN_0}{2K+1}\left(a_0 + \frac{\mu^2}{3kT}\right), \tag{2}$$

where μ is the electric dipole moment, a_0 the polarizability supposing the orientation fixed, and V the molar volume. In applying Onsager's formula to the powdered salts it will again be assumed that the effect of the air in the mixture

is negligible, and that N_0/V can be replaced by $n_1 = mN_0/Mv$. Equation (2) can then be written

$$\alpha' = \alpha_0 + \frac{\mu^2}{3kT} = \frac{(K-1)(2K+1)}{12\pi n_1 K}. \tag{3}$$

RESULTS

In Fig. 2 the quantity α' is plotted against the temperature for four different runs with hydrated chromium sulphate. The anomaly in the dielectric proper-

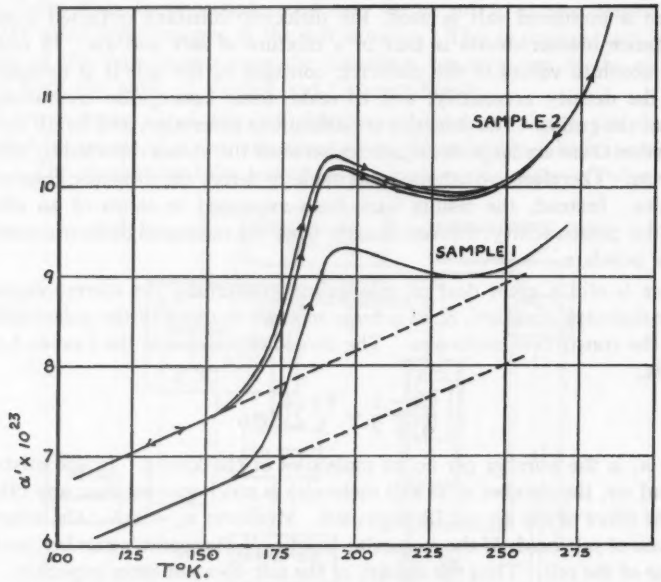

FIG. 2. Effective polarizability of hydrated chromium sulphate.

ties, with maximum close to the transition temperature of 195° K., shows clearly that the second-order transition does involve the orientation of electric dipoles. The nature of the transition will be discussed further below. Three runs made with the same filling of the cell are reasonably consistent, and there is no evidence of supercooling or temperature hysteresis, since at temperatures where they differ appreciably the curve for decreasing temperature lies between the two curves for increasing temperature. One run which was made with a different sample deviates considerably from the others in the absolute values of α', although the slope of the lower part of the curve, and the shape of the anomaly, are nearly the same. This can probably be attributed to the fact that the powder was less tightly packed, and that Onsager's formula applies properly only to a homogeneous material.

For the other salts studied there is no evidence of rotating dipoles, and Onsager's formula is less appropriate. However, in order to be consistent, and in order to compare the different salts, it is desirable to display the results in the same way. Therefore in Fig. 3 the quantity α' of Formula (3) is plotted against

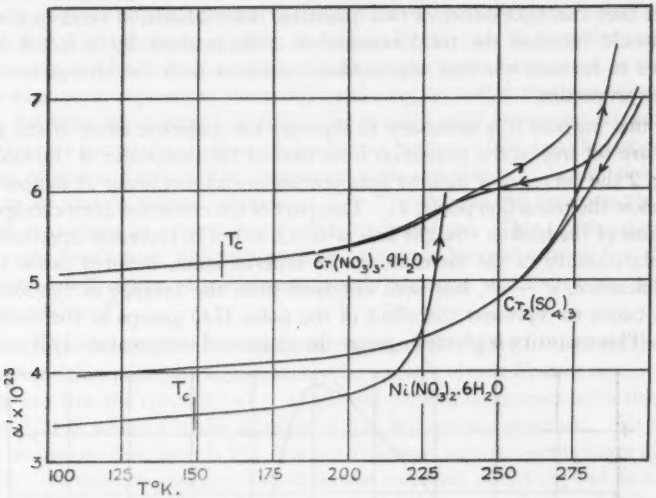

FIG. 3. Effective polarizability derived from sample runs with salts showing no dielectric anomalies.

the temperature for sample runs with chromium nitrate, nickel nitrate, and anhydrous chromium sulphate. There is no sign of any irregularity at the transition temperatures, T_c, of the nitrates, and so these transitions cannot involve electric dipoles. Again, different runs, made with different fillings of the cell, gave different absolute values for α'.

It will be noticed that for all the salts studied, including anhydrous chromium sulphate, the polarizabilities begin to increase very rapidly at about 230° to 250° K. This was accompanied by marked effects of dielectric absorption, as shown by the need for large changes in the compensating resistance R_3 in the bridge circuit, as well as by a marked dependence of the dielectric properties on frequency. Different fillings of the cell gave quite inconsistent results under these conditions, and in this temperature range no useful conclusions can be drawn from measurements made on the powdered salts.

NATURE OF THE TRANSITION IN CHROMIUM SULPHATE

Vasileff and Grayson-Smith (6) gave the value 1.41 R, where R is the gas constant, for the entropy of transition of $Cr_2(SO_4)_3 .18H_2O$. This is close to $R \ln 4 = 1.386\ R$, the agreement being as good as can be expected, since the experimental value depends on an extrapolation of the normal specific heat. According to Boltzmann's entropy theorem this indicates that the number of

configurations of the molecules is increased by a factor of 4 when the crystal passes into a disordered state above the transition temperature. The prominent change in the dielectric properties indicates that molecular groups having a large dipole moment are set free in the transition. A possibility which is consistent with the entropy change is that two of the 18 H_2O groups are released, so that they can take either of two quantized orientations, in random disorder. This would increase the total number of configurations by a factor 2^2. It remains to be seen whether this is also consistent with the change in the dielectric properties.

For this purpose it is necessary to separate the dielectric effect of the groups which are set free in the transition from that of the remainder of the molecule. In Fig. 2 the curves of α' against temperature are almost linear at temperatures well below the transition point, T_c. This part of the curve has been extrapolated by means of the broken straight line, which is taken to represent approximately the polarizability of the molecule in the ordered state, denoted below by α''. The difference, $\alpha' - \alpha''$, has been obtained from the average of the four runs and is taken to represent the effect of the polar H_2O groups in the disordered state. This quantity is plotted against the reciprocal temperature $1/T$ in Fig. 4.

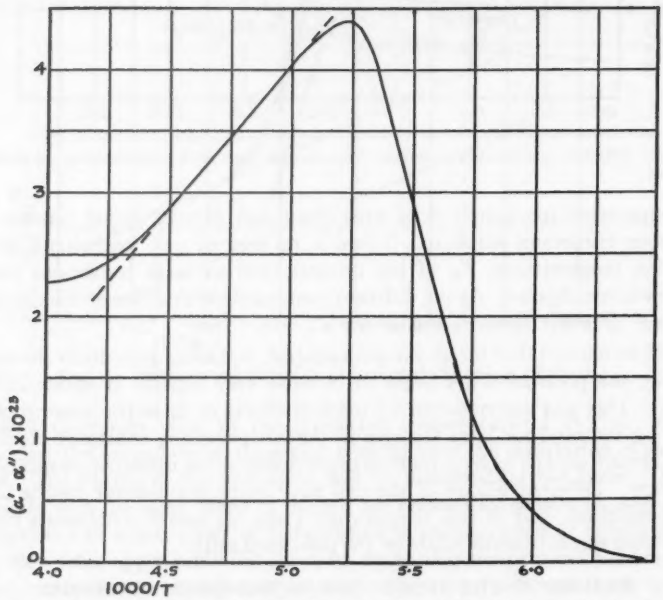

FIG. 4. Effect of the freeing of electric dipoles at the second-order transition in chromium sulphate.

Kirkwood (3) has given a theory of second-order transitions which involve rotations of electric dipoles and has given theoretical formulae for the excess entropy and for the quantity α' in Onsager's formula (Equation (3)). These

depend upon an "order parameter", s, which decreases from $s = 1$ at low temperatures (complete order) to $s = 0$ at $T = T_c$ (complete freedom of the rotating groups). Kirkwood's theory was developed for the case where the rotating groups can take any orientation at high temperatures and requires modification if the orientations are quantized. Apart from this modification, however, it gives only a qualitative description of the behavior of other crystals to which it ought to be applicable, and again in the case of chromium sulphate there is only qualitative agreement. In Fig. 4 the part of the curve to the right of the maximum represents the temperature region below the transition point, where the order parameter s is decreasing rapidly with increasing temperature, and where an increasing number of dipoles are being set free, but the dependence of s on temperature is not in agreement with the theory. Again, the theory predicts an abrupt change of slope at $T = T_c$, whereas the experiments give a smooth maximum at approximately this temperature. This is consistent with the specific heat curve, which shows that s does not become zero at $T = T_c$, and that complete freedom of orientation is only attained at a temperature several degrees above T_c.

Even if the details of the curve do not agree with theory in the neighborhood of the transition point, at temperatures far enough above T_c that nearly all the dipoles are free the quantity $\alpha' - \alpha''$ should depend on temperature through a term $q\mu^2/3kT$, where q is the number of free dipoles per molecule. To the left of the maximum the curve in Fig. 4 is nearly linear, up to a temperature of about 230° K. At this temperature the dielectric constant of the crystal as a whole begins to increase rapidly, as noted above, and the extrapolation of α'' becomes meaningless. The slope of the linear portion of the curve gives

$$q\mu^2/3k = 2.30 \times 10^{-20} \text{ c.g.s. electrostatic units.}$$

Assuming $q = 2$, as suggested above, this gives

$$\mu = 2.17 \times 10^{-18} \text{ e.s.u.}$$

for the electric moment of the dipole involved. The dipole moment of the H_2O molecule, derived from Stranathan's measurements of the dielectric constant of water vapor (5), is 1.831×10^{-18} e.s.u. The agreement is close enough to support the suggestion that the second-order transition in $Cr_2(SO_4)_3 \cdot 18H_2O$ is caused by the freeing of two of the 18 H_2O groups. For example, assuming $q = 1$ would give a value of the dipole moment so high as to be very unlikely. Moreover, a molecular group with a dipole moment much smaller than that of H_2O could not cause the large effect observed.

With regard to chromium nitrate and nickel nitrate, the absence of any change in polarizability at the transition points shows that the polar H_2O groups maintain their fixed orientations in the crystal. In both these materials, therefore, the second-order transitions are probably caused by rotations of the NO_3 groups.

ACKNOWLEDGMENT

The authors wish to express their thanks to the National Research Council for a grant-in-aid which made it possible to carry on the experimental measurements during the summer months.

REFERENCES

1. FRÖHLICH, H. Theory of dielectrics. Chapt. 2. Clarendon Press, Oxford. 1949.
2. JOHNSON, A. F. and GRAYSON-SMITH, H. Can. J. Research, A, 28: 229. 1950.
3. KIRKWOOD, J. G. J. Chem. Phys. 8: 205. 1950.
4. ONSAGER, L. J. Am. Chem. Soc. 58: 1486. 1936.
5. STRANATHAN, J. D. Phys. Rev. 48: 538. 1935.
6. VASILEFF, H. D. and GRAYSON-SMITH, H. Can. J. Research, A, 28: 367. 1950.

DESIGN AND USE OF A COINCIDENCE CIRCUIT OF SHORT RESOLVING TIME[1]

By R. E. Bell, R. L. Graham, and H. E. Petch[2]

ABSTRACT

A coincidence circuit for use with anthracene or stilbene scintillation counters is described. Its resolving time lies in the 10^{-9} sec. range, in agreement with theoretical estimates of the minimum possible resolving time with these scintillation counters. Examples are given of the use of the circuit in the measurement of short gamma-ray half-lives and, in conjunction with beta-ray spectrometers, in the determination of the details of radioactive disintegration schemes. The new results given in these examples are as follows:

Gamma energy, kev.	Emitted by	Parent activity	Half-life, sec.	Radiation shown to precede gamma-ray of first column
158	Hg^{199}	Au^{199}	$2.35 \pm 0.2 \times 10^{-9}$	297 ± 10 kev. beta
207	Hg^{199}	Au^{199}	$< 2 \times 10^{-10}$	250 ± 15 kev. beta
80	Xe^{131}	I^{131}	$4.8 \pm 2.0 \times 10^{-10}$	284 kev. gamma
364	Xe^{131}	I^{131}	$< 10^{-10}$	—
637	Xe^{131}	I^{131}	—	335 ± 15 kev. beta
722	Xe^{131}	I^{131}	—	250 ± 15 kev. beta
411	Hg^{198}	Au^{198}	$< 3 \times 10^{-11}$	—
238 (*F*-line)	ThC(Bi^{212})	ThB(Pb^{212})	$< 10^{-10}$	—

I. INTRODUCTION

The coincidence circuit to be described here is designed to exploit to the full the short resolving times possible with present day scintillation counters. The resolving times obtained are in the 10^{-9} sec. range.

The resolving time $2\tau_0$ of a coincidence circuit is equal to N_c/N_1N_2, where N_c is the chance coincidence rate measured when the two counters of the circuit are counting unrelated events at rates N_1 and N_2 respectively. The resolving time may also be measured by observing the prompt coincidence resolution curve of the coincidence circuit, i.e. the curve of coincidence rate as a function of the delay time artificially inserted in series with each counter in turn. We then have $2\tau_0$ equal to the effective width of the resolution curve (the area under the curve divided by its maximum height). For the usual shape of resolution curve found in practice, this figure is roughly equal to the full width at half maximum of the resolution curve. The resolving time is quoted as $2\tau_0$, the full effective width of the resolution curve, rather than the more usual half width, because of its fundamental nature mentioned above, and because the full width has a clear meaning when the resolution curve is not symmetrical. The symbol τ_0 is used to avoid confusion with τ, a mean life being measured with the coincidence circuit, but otherwise the notation agrees with older usage (7).

[1] *Manuscript received August 30, 1951.*
Contribution from the Nuclear Physics Branch, Chalk River Laboratory, Division of Atomic Energy Research of the National Research Council of Canada. Issued as N.R.C. No. 2640.
[2] *Now at Physics Department, University of British Columbia, Vancouver, B.C.*

The present coincidence circuit was first briefly described by Bell and Petch (4). It has been used in the measurement of gamma-ray half-lives near 10^{-9} sec. (3, 8) and, in conjunction with beta-ray spectrometers, in the investigation of disintegration schemes by coincidence techniques (8). Burcham (5) has used a similar circuit for the investigation of proton–gamma coincidences in (d,p) reactions. Petch and Johns (17) have used the circuit in an experiment on the angular correlation of the ThC″ gamma rays. McGowan (13) and Lundby (11) have recently described circuits which are similar in part. Bay (2) has recently published a coincidence circuit having a resolving time of a few times 10^{-10} sec. when used with the pulses from electron multipliers. It is shown in the next section, however, that the resolving time obtainable with anthracene and *trans*-stilbene scintillation counters is limited to the 10^{-9} sec. region by the limited speed of the phosphors, for electron energies in the region of 100 kev. We must therefore adopt the view that the scintillation counters, including their phosphors, form part of the coincidence circuit and cannot be logically separated from it.

II. ULTIMATE RESOLVING TIMES

The advantages of short coincidence resolving times are so obvious that it is of interest to estimate the ultimate resolving times possible with the presently available scintillation phosphors. We will restrict ourselves to the two best known organic phosphors, anthracene and *trans*-stilbene, excited by beta rays or secondary electrons produced by gamma rays. Suppose that each photomultiplier tube introduces no transit-time spread in its pulses, and that the coincidence circuit is of sufficient sensitivity to be actuated by the pulse due to a single photoelectron produced at the photocathode. It is clear that the resolving time of such an arrangement is limited only by the uncertainty in the time of ejection of the first photoelectron from the photocathode after the excitation of the phosphor by the particle being counted and that no refinements in the circuit can shorten the resolving time below this limit. Suppose the total number of photoelectrons produced during the pulse is R. Post and Schiff (18) have shown that the mean time delay for the appearance of the first photoelectron is

$$\bar{t} = \frac{\tau}{R}\left(1 + \frac{1}{R}\right), \quad (R \gg 1), \tag{1}$$

where τ is the mean life of the light flash from the phosphor, about 2.4×10^{-8} sec. for anthracene and 8×10^{-9} sec. for *trans*-stilbene. Thus the action of the phosphor is to give each counter a built-in mean life for the appearance of the first photoelectron equal to $\bar{t}$. In order to have a probability of 0.95 of observing the first photoelectron, we must wait a time roughly equal to $3\bar{t}$ after the excitation of the phosphor by the electron being counted. An otherwise perfect coincidence circuit having a resolving time $2\tau_0 = 6\bar{t}$ would then have a coincidence efficiency of $(0.95)^2 = 0.90$.* With RCA-1P21 photomultipliers and careful attention to light collection, we may expect $R \backsim E/3$ for anthracene and $R \backsim E/4$ for *trans*-stilbene, where E is the electron energy in kilo-

*That is, it would miss only 10% of the true coincidences occurring.

electron volts expended in the crystal. We can now write down the theoretical minimum resolving time for 90% coincidence efficiency, with the aid of Equation (1):

(2) $2\tau_0$ (theor. min.) $= 6\bar{t} \backsim 4.3 \times 10^{-7}\ (1 + 3/E)\ /E$ (anthracene),

(3) $2\tau_0$ (theor. min.) $= 6\bar{t} \backsim 1.9 \times 10^{-7}\ (1 + 4/E)\ /E$ (*trans*-stilbene).

It must be emphasized that Equations (2) and (3) give the limit due to the phosphor alone, assuming no contribution to the resolving time from transit-time spread in the photomultipliers or inadequacies in the rest of the coincidence circuit. Shorter resolving times can be achieved by allowing the coincidence efficiency to drop below 90%. Equations (2) and (3) are presented in graphical form in Fig. 1. The curves of Fig. 1 suggest minimum resolving times in the

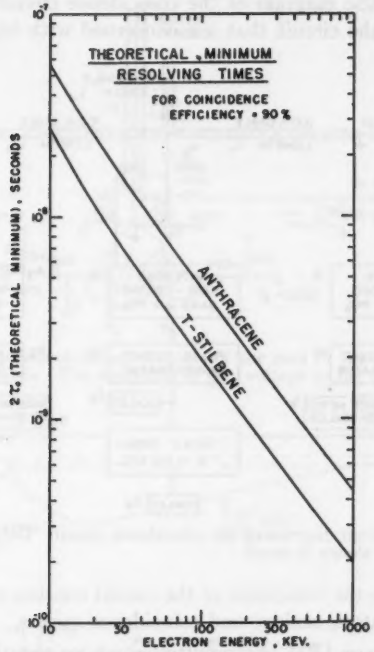

FIG. 1. Curves showing the estimated minimum possible resolving times for 90% coincidence efficiency for scintillation counters using anthracene and *trans*-stilbene phosphors and type 1P21 photomultipliers.

region of 100 kev. electron energy of 4.5×10^{-9} sec. for anthracene and 2×10^{-9} sec. for *trans*-stilbene. If the prompt coincidence resolution curve is plotted on a logarithmic scale, the sides of the resolution curve should have slopes numerically equal to $1/\bar{t}$. Thus the slope of the sides of the resolution curve should be such that the counting rate changes by a factor of two for an inserted delay of about one-ninth of these figures. Since the experimental performance of the present coincidence circuit agrees with these figures (Section IV), we may assume that

the circuit does in fact respond to the first photoelectron ejected from the photocathode, and that transit-time spread in the photomultipliers is not a serious effect at these resolving times.

It is clear from Fig. 1 that when measurements are being made near the limit of resolving time, care must be taken to see that the influence of the electron energy on the resolving time is correctly taken into account. It is also clear that the resolving times obtainable with electrons of unselected energy incident on the phosphor will be slightly longer than those shown in Fig. 1, which apply when electrons of a single energy are incident on each counter.

III. THE COINCIDENCE CIRCUIT

A complete schematic diagram of the coincidence circuit is given in Fig. 2. Only those parts of the circuit that are concerned with high-speed pulses are

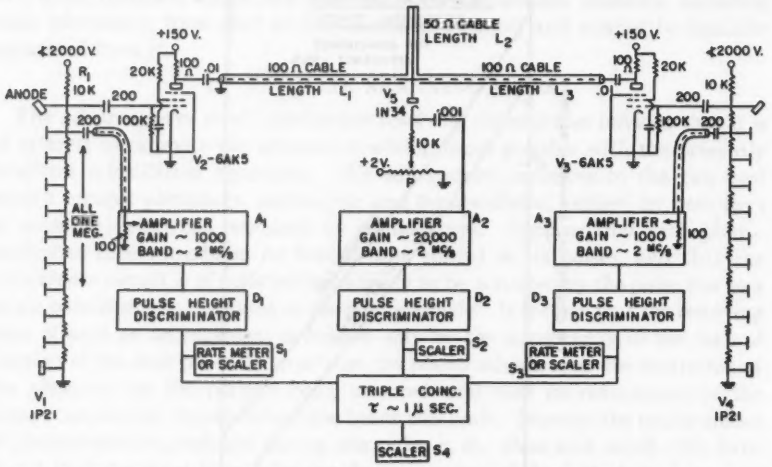

FIG. 2. Schematic circuit diagram of the coincidence circuit. Only those parts concerned with the short pulses are shown in detail.

shown in detail, since the remainder of the circuit consists only of conventional amplifiers, discriminators, scalers, and coincidence mixers.

In Fig. 2, V_1 is a type 1P21 photomultiplier whose negative anode pulses are impressed on the grid of V_2(6AK5). In order that the pulse produced by a single electron released from the photocathode of V_1 may be large enough ($\sim$ 3 v.) to cut off the standing plate current in V_2, the gain of the 1P21 must be about 2×10^8, if the capacity of V_1 anode plus V_2 grid is 10 $\mu\mu f$. The 1P21 tube must therefore be run at an applied voltage of at least 2000 v.; 2400 v. is customary. Some increase in the number of noise pulses relative to the number of true pulses is noticed at these voltages. 1P21 tubes are found to improve in this respect with continued application of high voltage. In most experiments either the pulse size is large enough so that the noise pulses may

be biased out, or the true counting rate is so high that very considerable noise counting rates are tolerable. In extreme cases it may be necessary to cool the photomultiplier. The fact that satisfactory coincidence results have been achieved with electron energies in the 40 to 70 kev. range (3, 8) without cooling suggests that cooling is necessary only for the lowest electron energies.

In Fig. 2, L_1 and L_3 are sections of coaxial cable of characteristic impedance $\backsim$ 100 ohms, terminated in 100 ohm resistors. These cable sections have no action other than that of serving as connecting leads and delaying the pulses from the individual counters. The action of the coincidence circuit can be understood from the simplified equivalent circuit of Fig. 3, where L_1 and L_3 have been omitted.

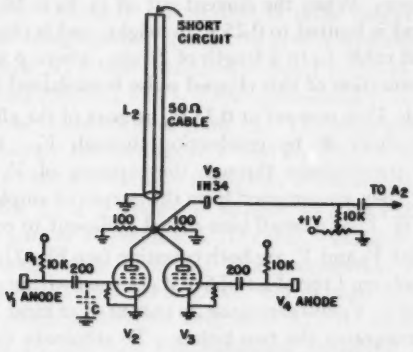

FIG. 3. Simplified equivalent circuit diagram of the part of the coincidence circuit concerned with the short pulses. The supplying of d-c. voltage to the plates of V_2 and V_3 has been ignored.

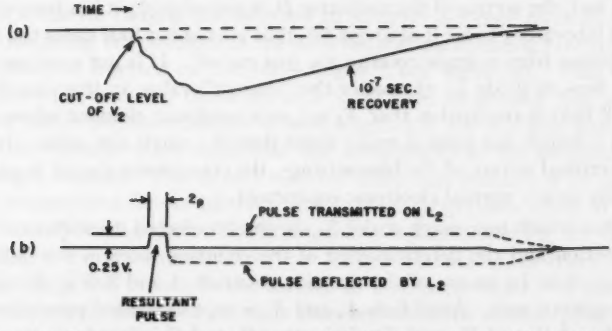

FIG. 4. (*a*) Voltage wave form at the anode of V_1 (or V_4) and grid of V_2 (or V_3) for a single counting event; (*b*) formation of the clipped and limited pulse at the plates of V_2 and V_3 in Fig. 3 or at the cable junction in Fig. 2. Neither of these wave forms has been observed other than by the over-all action of the coincidence circuit.

Suppose only V_1 is counting. A single counting event causes V_1 anode and V_2 grid to go negative with a pulse wave form of the type of Fig. 4(*a*). The

falling part of the pulse is the sum of a number of smaller pulses each due to the ejection of a single electron at the photocathode. The first of these smaller pulses is large enough to cut off the standing plate current in V_2. The smoothed out shape of the falling part of the pulse is

$$\left[1 - \exp\left(-\frac{t}{\tau}\right)\right],$$

where τ is again the phosphor decay time. The pulse recovers with the time constant R_1C, where $R_1 = 10$ kilohms and C is the capacity of V_1 anode plus V_2 grid, say 10 $\mu\mu$f. (Fig. 3). Thus R_1C is about 10^{-7} sec., and V_2 remains cut off for several times 10^{-7} sec., to guard against spurious pulses immediately following the main one. When the current cut off in V_2 is 10 ma., the pulse in the 25-ohm plate load is limited to 0.25 v. in height, and is clipped by reflection in the short-circuited cable L_2 to a length of $2p$ sec., where p is the transit time of cable L_2. The formation of this clipped pulse is explained by Fig. 4(*b*).

If the bias on diode V_5 is now set at 0.25 v., no part of the clipped and limited pulse can reach amplifier A_2 by conduction through V_5. In practice small pulses reach A_2 by transmission through the capacity of V_5 with widely distributed amplitude; they are removed from the output of amplifier A_2 by setting discriminator D_2 (Fig. 2) at a small bias E just sufficient to reject them all.

Supposing now that V_1 and V_4 are both counting (see Fig. 2), two pulses which occur within the resolving time of amplifier A_2 but *not* within $\pm 2p$ sec. of each other will not overlap at V_5 and can cause an output of at most $2E$ from amplifier A_2, since A_2 then integrates the two pulses. To eliminate this effect the bias of D_2 must be at least $2E$. When, however, the two pulses overlap at V_5, that part of the combined pulse greater than 0.25 v. is transmitted by conduction through V_5, and amplifier A_2 then has a large output, many times $2E$. Because of the last fact, the setting of discriminator D_2 is not critical; it has been customary in this laboratory to set it at about $4E$, that is at about four times the setting at which pulses from a single counter are just cut off. It is not even necessary to set the bias on diode V_5 at exactly the "correct" value, in this case 0.25 v., because all that is required is that V_5 act as a nonlinear element whose transmission of a double size pulse is many times that of a single size pulse. Because of the noncritical nature of the bias settings, the coincidence circuit is as stable in operation as any normal electronic equipment.

The pulses which now reach scaler S_2 are the unselected coincidences; pulse height selection and the determination of the counting rates in the individual counters are done by means of the two side channels, 1 and 3 (Fig. 2), and the triple coincidence unit. Amplifiers A_1 and A_3 accept untreated pulses from the last dynodes of V_1 and V_4, and discriminators D_1 and D_3 select from the output of A_1 and A_3 only those pulses larger than the desired amplitude. Much lower gains than those indicated in Fig. 2 would be acceptable for amplifiers A_1 and A_3 if the connecting cables from the photomultipliers were not terminated as shown. The termination is a convenience, however, because it makes the performance of the circuit independent of the length of these connecting cables.

The triple coincidence unit then selects out of all the coincidences appearing at S_2 only those that have been formed by pulses whose amplitudes lie within the ranges selected by D_1 and D_3. This method of postcoincidence pulse height selection avoids the difficulties of pulse height selection where very high-speed pulses are to be transmitted. By an obvious extension of the circuit of Fig. 2, the simple discriminators D_1 and D_3 may be replaced by single-channel or multichannel pulse height analyzers, for scintillation spectrometry in conjunction with the coincidence circuit.

The triple coincidence unit may be of any conventional type. In the present equipment its resolving time is about 1 μsec. It is easily seen that chance triple coincidences cannot contribute appreciably to the chance coincidence rate of the over-all system. Owing to this system of pulse height selection, it is impossible for a pulse from a single counter to give a spurious coincidence pulse.

The resolving time $2\tau_0$ of the coincidence circuit is chiefly determined by the length of the clipping cable L_2, and is given approximately by $2\tau_0 = 4p$ where p, as before, is the pulse transit time of cable L_2. The observed value of $2\tau_0$ is also somewhat influenced by the settings of P and D_2 (Fig. 2) and by the electron energy dissipated in the scintillation phosphor, as suggested previously.

In practice, cable L_2 is a length of 50-ohm RG/8U cable, and L_1 and L_3 consist at least in part of 100-ohm RG/7U cable. Variable delay of the pulses is introduced by lengthening L_1 or L_3. The propagation velocity for RG/8U cable is $0.67c$, and for RG/7U cable $0.83c$. For manual operation, delay times can be varied by plugging different lengths of RG/7U cable in and out of the circuit in positions L_1 and L_3. For automatic operation, two variable-delay units have been used in this laboratory. The first is a pair of motor-driven wafer-type rotary switches by which graded lengths of RG/7U cable are switched in and out of the circuit, giving a total range of delay of 2.7×10^{-8} sec. Provision is made for causing the switches to advance at predetermined intervals, and for recording the positions of the switches and the counting rate appropriate to the delay so inserted on Esterline–Angus 1 ma. recorders. This device is satisfactory for resolving times $2\tau_0$ of 5×10^{-9} sec. or greater, but at shorter resolving times the uncertainties in lead length and the variable loss of signal due to the variable stray capacities and inductances of the switches become increasingly evident. The second variable delay unit consists of an air-spaced coaxial transmission line of 100 ohm characteristic impedance made in the form of a helix of $4\frac{1}{2}$ in. diameter. The arrangement is sketched in Fig. 5. The outer conductor of the transmission line was formed by cutting, in effect, a square cross section screw thread in the surface of a brass cylinder. A hard drawn copper wire supported on lucite posts then acts as center conductor of a coaxial transmission line in a square shield with one side open. A small carriage with a contact pressing on the center conductor contains L_2, V_5, and the first stage of amplifier A_2 (see Fig. 2). It is driven along the screw thread as the drum is turned, so that delay is taken out of one side of the circuit and inserted into the other. The pulses from the photomultipliers are led in through the end bearings of the rotating drum. Switches and index contacts are provided so that the

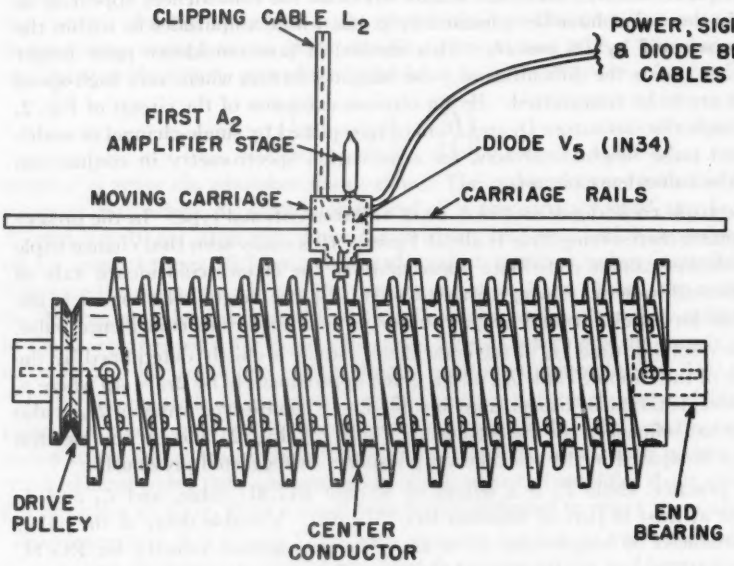

FIG. 5. Sketch explaining the principle of the helical delay unit. The drawing is not to scale, and most of the mechanical and electrical drive details have been omitted. The short pulses enter and leave through the end bearings.

drum may be stopped automatically at any one of 57 different delay positions covering a total range of 6.4×10^{-8} sec., the smallest delay step being 4×10^{-10} sec. This unit, though complex mechanically, is electrically very satisfactory, and shows almost no sign of variable loss in pulse size as the delay is changed. Either of the two automatic delay units converts the coincidence circuit into a single-channel "time-sorter" of channel width equal to $2\tau_0$. All coincidence resolution curves shown in this paper were recorded with one or other of these two delay units, often over periods of many hours.

The scintillation counters for the circuit are fully mobile, being mounted on small metal boxes which contain the limiting pentodes (V_2 or V_3, Fig. 2). Cables connect the small boxes to the appropriate power supplies and to the delay unit and amplifier chasses. All leads for the high speed parts of the circuit are kept short and direct, since a band width of some hundreds of megacycles is required to preserve the form of these short pulses.

IV. RESULTS

(a) *Performance of the Circuit*

The performance of the coincidence circuit is best judged by measuring its resolution curve under known conditions. Typical resolution curves for gamma–gamma coincidences with anthracene phosphor crystals were shown in a previous publication (4) indicating a resolving time $2\tau_0$ of about 5×10^{-9} sec. Further curves will be shown here indicating how the use of the faster *trans*-stilbene phosphor enables the resolving time to be shortened.

Resolution curves are measured by observing the coincidence counting rate as a function of the delay time inserted in series with one of the counters. (Delay inserted in series with one of the counters is of course the equivalent of negative delay inserted in series with the other.) The resolution curve of Fig. 6 was

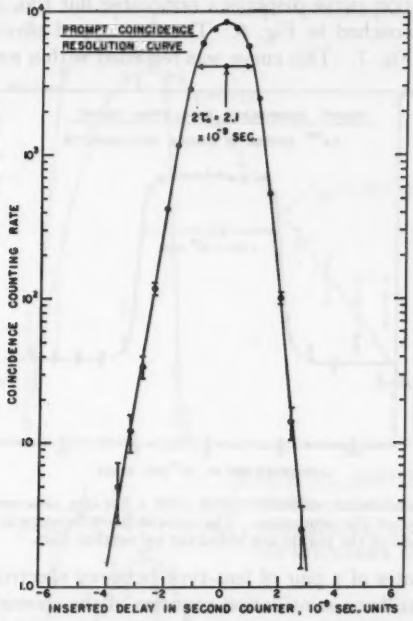

FIG. 6. Prompt coincidence resolution curve observed when hard beta rays were sent through the thin *trans*-stilbene crystal of the first counter into the *trans*-stilbene crystal of the second counter. No chance coincidences have had to be subtracted, since none occur in such an arrangement. The standard deviations of the points are indicated by vertical bars.

recorded for the case when hard beta particles from a radioactive source were passed through the thin *trans*-stilbene crystal of one counter into the *trans*-stilbene crystal of the second counter. The pulse size from the thin crystal was equal to that produced by electrons of about 100 kev. The chief advantage of this method is that the chance coincidence rate is zero, and the true coincidence rate may be followed down for several decades from its peak without any

interference from chance events. The curve of Fig. 6 shows a resolving time $2\tau_0 = 2.1 \times 10^{-9}$ sec. The slopes of the sides of the curve are such that the coincidence rate falls by a factor of two in 2.4×10^{-10} sec. (left side of curve) or 1.6×10^{-10} sec. (right side of curve). The slight asymmetry of the curve of Fig. 6 is attributable to the fact that the pulses from the first counter, whose crystal is traversed by the beta particles, are more nearly uniform in size than those from the second counter. The resolution curve of Fig. 6 was recorded with a clipping cable L_2 of air-equivalent length 15 cm. and hence a transit time p of 5×10^{-10} sec. The expected resolving time $2\tau_0$ is therefore $4p = 2 \times 10^{-9}$ sec., in good agreement with the observed value. The coincidence efficiency for this case was about 85%.

When the resolving time is made long enough so that the coincidence efficiency is 100%, the resolution curve displays a noticeable flat top, a condition which was just being approached in Fig. 6. The flat-topped effect is shown in the resolution curve of Fig. 7. This curve was recorded with a source of Au^{198} on a

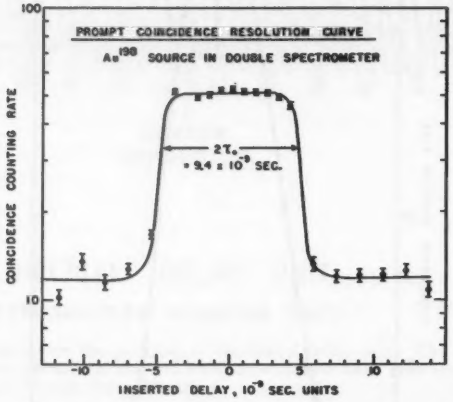

FIG. 7. Prompt coincidence resolution curve with a flat top, observed when the resolving time is lengthened beyond the minimum. The coincidence efficiency in this case was 100%. The standard deviations of the points are indicated by vertical bars.

thin film at the center of a pair of lens-type beta-ray spectrometers placed end to end. The two stilbene scintillation counters of the coincidence circuit acted as detectors for the two spectrometers. One spectrometer was set to focus the K-conversion line of the 411 kev. gamma ray which follows the beta rays of Au^{198}, and the other spectrometer was set on the beta continuum beside the line. The 411 kev. gamma ray is shown later in this paper to have a half-life near 10^{-11} sec., so that Fig. 7 shows a curve which is effectively a prompt resolution curve. The coincidence efficiency could now be checked as follows: From Fig. 7 and a knowledge of the counting rates in the two counters (scalers S_1 and S_3, Fig. 2) the strength of the Au^{198} source could be calculated. The result was $(1.260 \pm 0.019) \times 10^7$ disintegrations per second, assuming a coincidence efficiency of 100%. The same source was allowed to decay for 14 days and then

measured by conventional beta–gamma coincidence methods by R. C. Hawkings. The latter result, corrected for decay, was $(1.268 \pm 0.028) \times 10^7$ disintegrations per second. The agreement of the two results shows that the assumption of 100% coincidence efficiency is correct to within the combined standard deviation of the measurements. The coincidence circuit may thus be used for assaying source strengths by the coincidence method; owing to the short resolving time available much larger sources may be measured than is usual in this technique, with correspondingly reduced time of measurement.

(*b*) *Measurement of Short Half-lives*

The delayed-coincidence method has been used many times for the measurement of short half-lives (7). The introduction of coincidence circuits with re-

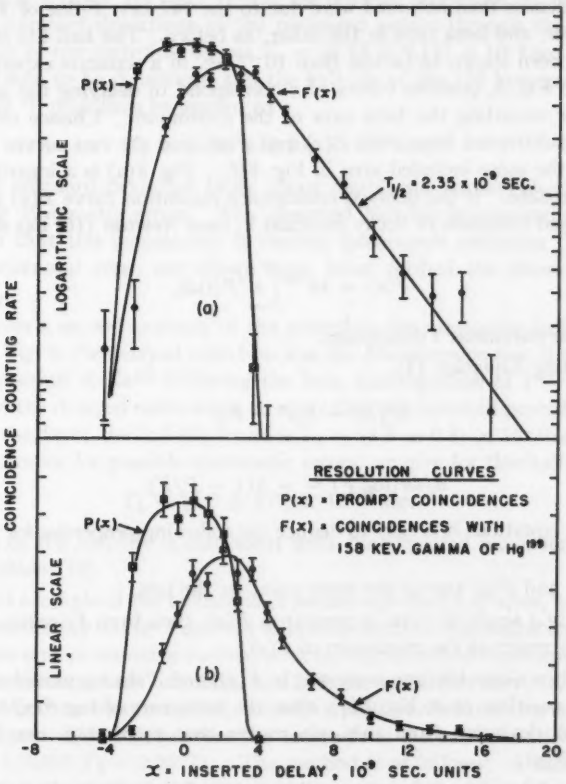

FIG. 8. Delayed coincidence resolution curve $F(x)$ for the 158 kev. gamma ray of Hg^{199}, with a prompt curve $P(x)$ for comparison; (*a*) logarithmic scale, (*b*) linear scale. The half-life of the gamma ray can be measured by the slope of the right-hand part of log $F(x)$ in (*a*) or by the shift of the centroid of $F(x)$ to the right in (*b*). The result is $T_{\frac{1}{2}} = (2.35 \pm 0.20) \times 10^{-9}$ sec. The standard deviations of the points are indicated by vertical bars.

solving times in the 10^{-9} sec. range and the theoretical analyses of Newton (16) and Bay (1) have combined to permit the measurement by this method of half-lives as short as a few times 10^{-10} sec. The method may be understood by reference to Fig. 8, which gives the results of a delayed-coincidence experiment performed on the L-conversion line of the 158 kev. gamma ray emitted by Hg^{199} following the beta disintegration of Au^{199}. The experiment was carried out in the pair of beta-ray spectrometers mentioned above. One spectrometer was held focused on the 145 kev. L-conversion peak of the 158 kev. gamma ray, while the second was focused on beta rays of the continuum.

In Fig. 8, $F(x)$ is the delayed coincidence resolution curve observed under these conditions, where x is the inserted time delay. The prompt resolution curve $P(x)$ was obtained by replacing the source of Au^{199} by a source of ThB. The coincidences then obtained were due to the 147 kev. F-line of ThB in one spectrometer, and beta rays in the other, as before. The half-life of the ThB F-line had been shown to be less than 10^{-10} sec. in a separate experiment (see below). In Fig. 8, positive values of x correspond to delaying the pulses from the counter recording the beta rays of the continuum. Chance coincidences have been subtracted from both $P(x)$ and $F(x)$, and the two curves have been plotted to the same included area in Fig. 8(b). Fig. 8(a) is a logarithmic plot of the same data. If the delayed coincidence resolution curve $F(x)$ is due to a single delayed radiation of decay constant λ, then Newton (16) has shown that

$$F(x) = \lambda e^{-\lambda x}\int_{-\infty}^{x} e^{\lambda t}P(t)dt, \tag{4}$$

in which the parameter t disappears.

Differentiating Equation (4),

$$dF/dx = \lambda(P - F), \tag{5}$$

and

$$d/dx(\log F) = -\lambda(1 - P/F). \tag{6}$$

From these equations it is easy to deduce the following properties for the curves of Fig. 8:

(a) $F(x)$ and $P(x)$ rise at the same point on the left,

(b) if only a single lifetime is present in $F(x)$, then from Equation (5), $F(x)$ and $P(x)$ intersect at the maximum of $F(x)$,

(c) if only a single lifetime is present in $F(x)$, and if the curvature of log $P(x)$ is nowhere positive, as in Fig. 8(a), then the curvature of log $F(x)$ is nowhere positive. A deviation from this rule means that more than one lifetime is present in $F(x)$,

(d) at positive delays great enough so that $F(x) \gg P(x)$, $F(x)$ falls off as $\exp(-\lambda x)$, or from Equation (6),

$$d/dx(\log F) = -\lambda, \; F \gg P, \tag{7}$$

(*e*) the centroid of $F(x)$ is displaced positively along the x-axis from the centroid of $P(x)$ by an amount $\tau = 1/\lambda$, the mean life of the delayed radiation, as first shown by Bay (1).

Of these properties of coincidence resolution curves, (*a*), (*b*), and (*c*) are useful checks on the consistency of a delayed-coincidence experiment, and (*d*) and (*e*) may be used to evaluate the mean life that is being measured. All the properties are illustrated in Fig. 8(*a*) and (*b*). In evaluating the mean life, (*d*) is most useful when τ is long compared with the resolving time $2\tau_0$, and (*e*) is essential when τ is short compared with $2\tau_0$. Fig. 8 represents an intermediate case where both (*d*) and (*e*) can be applied. (As usual in experimental work, the results are quoted as half-lives rather than mean lives.) Using Equation (7) and fitting by least squares all points at positive delays greater than 4×10^{-9} sec., we get $T_{\frac{1}{2}} = (2.26 \pm 0.12) \times 10^{-9}$ sec. Using the centroid procedure (*e*), carrying the statistical standard deviations of the measured points through the numerical computation of the centroids, we get $T_{\frac{1}{2}} = (2.43 \pm 0.12) \times 10^{-9}$ sec. The two results are seen to be consistent, and the half-life of the 158 kev. gamma transition in Hg ^{199} may thus be quoted as

$$T_{\frac{1}{2}} = (2.35 \pm 0.20) \times 10^{-9} \text{ sec.},$$

the quoted standard deviation being about double the statistical one to allow for possible systematic errors. The observed half-life is consistent with the assumption that this gamma ray is electric quadrupole radiation (19). Two other experimental runs, not shown here, have verified the above measured value of $T_{\frac{1}{2}}$.

Fig. 9 gives a second example of the procedure for measuring half-lives. In the case of Fig. 9, the delayed radiation was the K-conversion line of the 80 kev. gamma transition of Xe^{131} following the beta disintegration of I^{131}. Here the mean life of the delayed radiation is so short that the centroid procedure makes best use of the data; the half-life found is $T_{\frac{1}{2}} = (4.8 \pm 0.8) \times 10^{-10}$ sec. Again making allowance for possible systematic errors, we give for this half-life

$$T_{\frac{1}{2}} = (4.8 \pm 2.0) \times 10^{-10} \text{ sec.}$$

This value of the half-life is consistent with the expected value for magnetic dipole radiation (19).

As a third example of the technique of measuring short half-lives, we take the 411 kev. gamma ray of Hg^{198} emitted following the beta disintegration of Au^{198}. For this case a more sensitive method of detecting short half-lives was used, the so-called "self-comparison" method, in which the resolution curve for the Hg^{198} conversion electrons was compared, not with a prompt resolution curve, but with its own inverse. The shift between the centroids of the two resolution curves is then $2\tau = 2/0.693\ T_{\frac{1}{2}} = 2.89\ T_{\frac{1}{2}}$. The method is as follows: A thin source of Au^{198} on a thin aluminum foil is placed at the center of the pair of lens beta-ray spectrometers mentioned previously. Let the two spectrometers be called the north and the south spectrometer. For the first run, the north spectrometer is focused on the beta-ray continuum just below the K-conversion line of the 411

kev. gamma ray, and the south spectrometer is focused on the line. For the second run, both spectrometers have their currents increased by the same small percentage, so that the north spectrometer is now on the line, and the south spectrometer is on the beta-ray continuum just above the K-conversion line. Thus all effects on the coincidence resolution curve arising from changes between runs in the electron energy or momentum are small and in the same direction for both spectrometers. At the same time, each counter is counting only electrons of a single energy at any one time, and this is the condition for the best resolution curve.

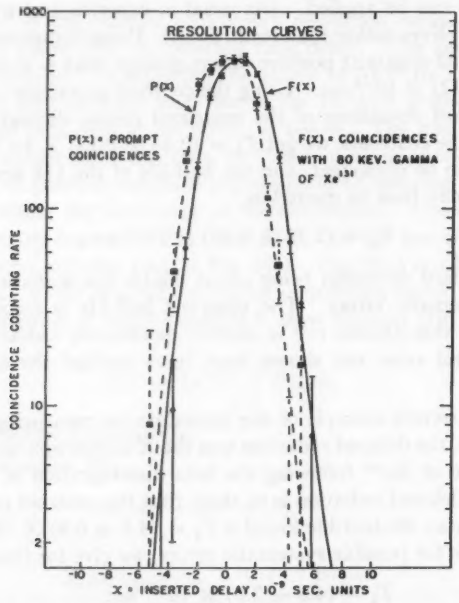

FIG. 9. Delayed coincidence resolution curve $F(x)$ for the 80 kev. gamma ray of Xe^{131}, with a prompt curve $P(x)$ for comparison. The half-life of the gamma ray is measured by the shift of the centroid of $F(x)$ to the right. The result is $T_{\frac{1}{2}} = (4.8 \pm 2.0) \times 10^{-10}$ sec. The standard deviations of the points are indicated by vertical bars.

The pair of resolution curves found in this experiment is shown in Fig. 10. The curve resulting from the second run (lower curve in Fig. 10) should lie to the right of the curve resulting from the first run (upper curve in Fig. 10) by 2.89 $T_{\frac{1}{2}}$. It is obvious from Fig. 10 that the shift is very small. Applying the centroid calculation to the points of Fig. 10 we find for the 411 kev. gamma ray of Hg^{198}

$$T_{\frac{1}{2}} = (1.0 \pm 1.7) \times 10^{-11} \text{ sec.}$$

Since the standard deviation is greater than the measured value, this result means that the half-life of the 411 kev. gamma ray is less than, say, 3×10^{-11}

sec. Moon (15) has shown by the resonant nuclear scattering of gamma rays that the half-life of this gamma ray cannot be much less than about 10^{-11} sec. The present upper limit combined with Moon's lower limit shows that the half-life of this gamma ray lies in the region of 10^{-11} sec. Previous measurements by the delayed coincidence method fell into two groups, one agreeing with the present result by showing a half-life less than about 3×10^{-9} sec. (4, 6) and the other showing a half-life near 3×10^{-8} sec. (14, 9, 12), a figure exceeding that derived here by a factor of 10^2. The cause of the discrepancy probably lies

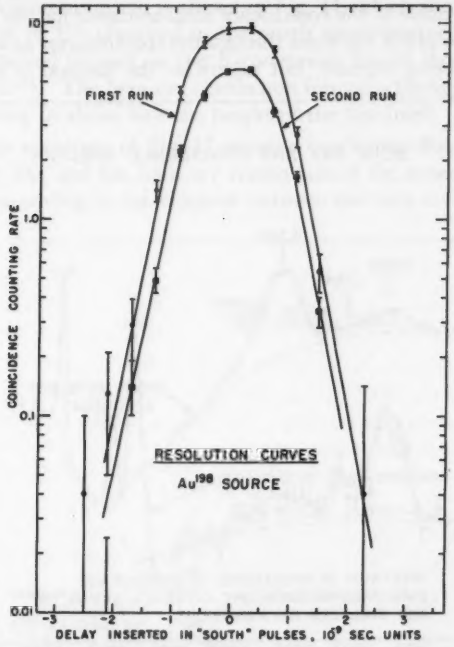

FIG. 10. Self-comparison resolution curves for the 411 kev. gamma ray of Hg^{198}. The delayed resolution curve marked "second run" is the inverse of the curve marked "first run", and should lie to the right of the "first run" curve by 2.89 $T_{\frac{1}{2}}$. The result for $T_{\frac{1}{2}}$ is $(1.0 \pm 1.7) \times 10^{-11}$ sec. The standard deviations of the points are indicated by vertical bars.

either in defects in some earlier delayed coincidence techniques, or in lack of identification of the radiation causing the observed half-life in the earlier experiments. A half-life of the order of 10^{-11} sec. is consistent with the theoretical value for electric quadrupole radiation (19).

Similar experiments by the self-comparison method have shown that the half-lives of the *F*-line of ThB and of the 364 kev. gamma ray of Xe^{131} following the beta disintegration of I^{131} are both less than 10^{-10} sec., and that the half-life of the 207 kev. gamma ray of Hg^{199} following the beta disintegration of Au^{199} is less than 2×10^{-10} sec.

(c) *Coincidence Measurements in the Beta-ray Spectrometer*

The coincidence circuit has been used in the study of radioactive disintegration schemes in conjunction with the pair of beta-ray spectrometers mentioned previously. In most cases a radioactive source is mounted on a thin foil at the center of the pair of lens spectrometers; with the two counters of the coincidence circuit acting as detectors for the two spectrometers, time correlations between different radiations focused in the two spectrometers can then be investigated. In addition, of course, ordinary beta-ray spectroscopy can be carried on. Allowance must be made in the coincidence measurements for the finite delay due to the transit time of the electrons through the spectrometer as a function of the electron energy being focused, but otherwise the method is straightforward.

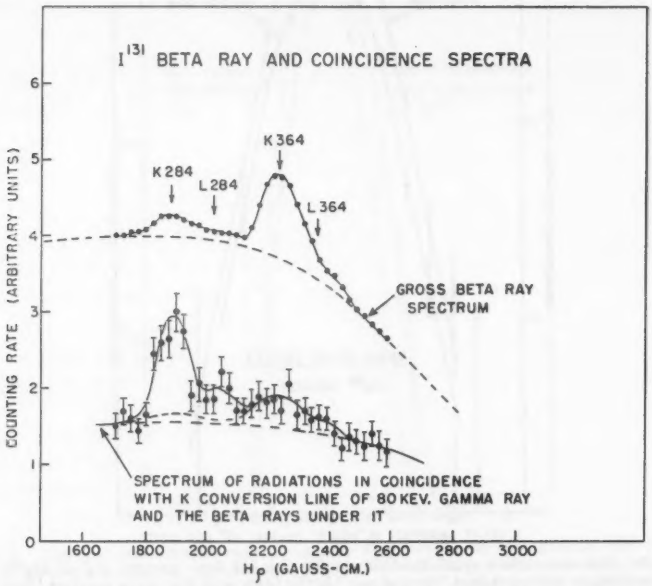

FIG. 11. Section of the beta-ray spectrum of I^{131} in one spectrometer plotted together with the spectrum of coincidences observed when the other spectrometer is held focused on the K-conversion line of the 80 kev. gamma ray of Xe^{131}, and, of necessity, on the beta-ray continuum forming a background under it. The background continuum amounted to about half the height of the $K80$ line. This figure shows that the 80 kev. gamma ray is in cascade with the 284 kev. gamma ray but not with the 364 kev. gamma ray. The standard deviations of the points are indicated by vertical bars.

The source strength used is typically of the order of 0.2 mc. Each spectrometer has a resolution in momentum of the order of 4% and a transmission of about 0.4%. The true coincidence counting rates obtained vary from less than 0.1 to more than 100 per minute. Automatic current-switching is used for the spectrometers, with automatic recording of results. One of the motor-driven delay units can be connected into the circuit so as to record successively true-plus-

chance coincidences and chance-alone coincidences for each setting of the spectrometer coil current.

As a first example of the use of coincidence beta-ray spectroscopy, we take the prominent 80, 284, and 364 kev. gamma transitions in Xe^{131} following the beta disintegration of I^{131}. Lind *et al.* (10) have shown that the sum of the energies of the 80 and 284 kev. gammas is equal to that of the 364 kev. gamma within one part in 4500. There is thus a strong presumption that the 80 and 284 kev. gamma rays form a cascade pair in parallel with the 364 kev. gamma ray. The direct proof of this presumption is shown in Fig. 11, which shows a part of the beta-ray spectrum of I^{131} observed in the north spectrometer while the south spectrometer remained focused on the K-conversion line of the 80 kev. gamma ray (the "$K80$ line"). The beta-ray continuum formed a background under the $K80$ line amounting to about half the height of the line itself.

The coincidence spectrum of Fig. 11 shows a continuum due to coincidences between the $K80$ line and the beta-ray continuum of the gross spectrum, plus small peaks corresponding to coincidences between the beta rays under the $K80$

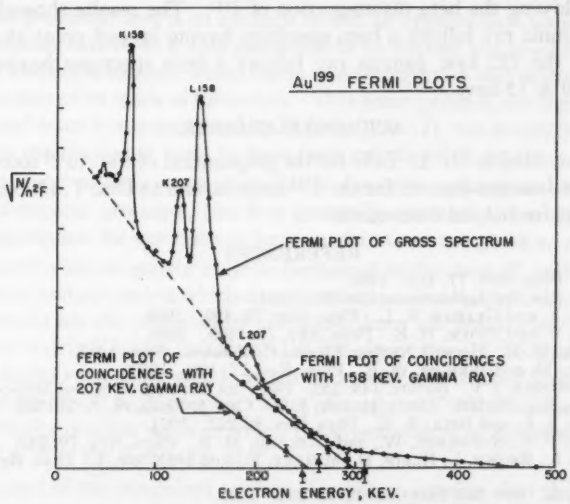

FIG. 12. Fermi plots of the gross beta-ray spectrum of Au^{199} in one spectrometer and of the spectrum of coincidences observed when the other spectrometer is held focused in turn on conversion lines of the 158 and 207 kev. gamma rays of Hg^{199}. This figure shows that the 158 kev. gamma ray follows a beta spectrum having its end point at 297 ± 10 kev., and that the 207 kev. gamma ray follows a beta spectrum having its end point at 250 ± 15 kev. The standard deviations of the points are indicated by vertical bars.

line and the conversion lines of the 284 and 364 kev. gamma rays, plus relatively large peaks due to coincidences between the $K80$ line and the K and L conversion lines of the 284 kev. gamma ray. The conclusion about the arrangement of the Xe^{131} gamma rays follows from this analysis of Fig. 11.

A second example is taken from the spectrum of Au^{199}. In this case the L-conversion line of the 158 kev. gamma ray of Hg^{199} was held focused in the north spectrometer while the south spectrometer was scanned over the Au^{199} beta spectrum. The results are presented in the form of Fermi plots of the gross beta-ray spectrum and the coincidence spectrum in Fig. 12. It is seen that, where the Fermi plot of the gross spectrum has an extrapolated end point of 460 ± 10 kev. and appears to consist of two or possibly three components, the coincidence Fermi plot has an end point of 297 ± 10 kev. The 158 kev. gamma ray therefore follows a beta spectrum having an end point of 297 ± 10 kev. As shown in Fig. 12, a similar experiment for the 207 kev. gamma ray has demonstrated that the 207 kev. gamma ray of Hg^{199} follows a beta spectrum of Au^{199} having an end point of 250 ± 15 kev., and hence that the gross beta-ray spectrum does consist of at least three components. This second result particularly shows the value of the above coincidence procedure, since it is difficult to pick out the 250 kev. end point on the gross beta-ray Fermi plot, even if the unwarranted assumption is made that all the beta-ray components have the allowed shape.

A similar experiment has been performed on the 637 and 722 kev. gamma rays of Xe^{131} following the beta disintegration of I^{131}. The results showed that the 637 kev. gamma ray follows a beta spectrum having its end point at 335 ± 15 kev., while the 722 kev. gamma ray follows a beta spectrum having its end point at 250 ± 15 kev.

ACKNOWLEDGMENTS

We are indebted to Dr. L. Yaffe for the preparation of the Au^{199} source, to the Chalk River Isotopes Branch for the I^{131} sources, and to Drs. T. D. Newton and L. G. Elliott for helpful discussions.

REFERENCES

1. BAY, Z. Phys. Rev. 77: 419. 1950.
2. BAY, Z. Rev. Sci. Instruments, 22: 397. 1951.
3. BELL, R. E. and GRAHAM, R. L. Phys. Rev. 78: 490. 1950.
4. BELL, R. E. and PETCH, H. E. Phys. Rev. 76: 1409. 1949.
5. BURCHAM, W. E. Harwell Nuclear Physics Conference. Sept. 1950.
6. DEUTSCH, M. and WRIGHT, W. E. Phys. Rev. 77: 139. 1950.
7. e.g. DUNWORTH, J. V. Nature, 144: 152. 1939; DE BENEDETTI, S. and MCGOWAN, F. K. Phys. Rev. 70: 569. 1946; JELLEY, J. V. Can. J. Research, A, 26: 255. 1948.
8. GRAHAM, R. L. and BELL, R. E. Phys. Rev. 83:222. 1951.
9. JASTRAM, P. S., KONNEKER, W., and CLELAND, M. R. Phys. Rev. 79: 243. 1950.
10. LIND, D. A., BROWN, J., KLEIN, D., MULLER, D., and DUMOND, J. Phys. Rev. 75: 1544. 1949.
11. LUNDBY, A. Rev. Sci. Instruments, 22: 324. 1951.
12. MACINTYRE, W. J. Phys. Rev. 76: 312. 1949.
13. MCGOWAN, F. K. Oak Ridge Progress Report. Sept. 1950 (ORNL-865).
14. MEYER, K. P., BALDINGER, E., and HUBER, P. Helv. Phys. Acta, 23: 517. 1950.
15. MOON, P. B. Proc. Phys. Soc. (London), A, 64: 76. 1951.
16. NEWTON, T. D. Phys. Rev. 78: 490. 1950.
17. PETCH, H. E. and JOHNS, M. W. Phys. Rev. 80: 478. 1950.
18. POST, R. F. and SCHIFF, L. Phys. Rev. 80: 1113. 1950.
19. WEISSKOPF, V. F. and BLATT, J. Chapter from forthcoming book on nuclear theory, privately circulated.

THE COLLISION MATRIX FOR THE COMPOUND NUCLEUS[1]

By T. D. Newton

ABSTRACT

The nuclear collision matrix is calculated under the restrictions of the compound nucleus hypothesis. The matrix so found contains only one undetermined function, $\tau(E)$, a phase shift caused by nuclear interaction. Cross sections are evaluated in terms of $\tau(E)$, a set of constants, and functions determined by the non-nuclear potentials.

1. INTRODUCTION

The most convenient tool for the discussion of nuclear reactions among many particles is the so-called collision matrix developed particularly by E. P. Wigner and his collaborators (6 and 7). This gives the amplitudes of the outgoing waves of scattered particles and reaction products when the amplitudes of the ingoing waves are assigned. Unfortunately the general form of this matrix is somewhat complicated.

The work presented here is an attempt to find an approximate simplified form for the collision matrix **U**. The procedure is to impose the condition that a compound nucleus is formed, whose mode of disintegration and internal density are independent of its mode of formation. This assumption is just the one which has been used since it was suggested by Bohr in 1936 (2), but its implications for the collision matrix do not seem to have been calculated in detail.

The compound nucleus concept probably does not provide an accurate description of nuclear processes, but it is generally assumed to be a reasonably good approximation for reactions in large nuclei at not too great an energy.

The general collision matrix may be factorized in the form $\mathbf{U} = \Omega\mathbf{V}\Omega$, where Ω is a diagonal unitary matrix which contains the so-called "potential scattering", i.e. its elements are the phase-shifts calculable from the non-nuclear potential. The reduced matrix **V** is also unitary and so has eigenvalues of the form $\exp(2i\tau)$, where the angles τ are phase shifts caused by specifically nuclear forces. In the general case there will be N possible values of τ, where N is the number of different energetically possible ways in which the given system of nucleons may disintegrate, i.e. N is the dimension of **U**.

The first part of the compound nucleus assumption requires that there be only two distinct eigenvalues of **V**; the value 1, which is $(N-1)$ fold degenerate, and one nontrivial, nondegenerate value $\exp(2i\tau)$; i.e. that only one specifically nuclear phase shift arises from a given compound nucleus.

The assumption that the interior density is the same for any choice of incoming currents then establishes the form of the potential scattering terms and of the penetration factors. The final result is that all cross sections for processes which go through a given compound nucleus may be expressed in terms of the

[1] *Manuscript received August 16, 1951.*
Contribution from the Atomic Energy Project, National Research Council of Canada, Chalk River, Ontario. Issued as N.R.C. No. 2639.

energy function $\tau(E)$, a finite set of constants (one for each mode of splitting), and quantities calculable from the external potential.

The usual criterion for the formation of a compound nucleus is that the branching ratios for reaction products be independent of the incoming currents. There is another criterion which does not seem to have been noticed previously; for each particular compound nucleus there is a value of the energy between each pair of resonances at which the cross section for every reaction (except scattering) is zero. If the existence of this zero could be shown for some system, it would be good evidence of the validity of the compound nucleus concept for that system. However the observation would be difficult because most experiments may form several different compound nuclei; i.e. several different spin values.

The above results are obtained in Section 3 and discussed in Section 4. The scattering and absorption cross sections are, of course, among those obtained from this development, and they bear a close resemblance to those obtained by Feshbach, Peaslee, and Weisskopf (3). On the other hand the form imposed on Wigner's general matrix by the compound nucleus hypothesis is found in the appendix. Hence the present paper demonstrates the relation between these two treatments of the problem.

The work is restricted to disintegrations into just two fragments. Non-relativistic equations are used throughout, so the results do not apply, in principle, to γ- and β-processes.

There is one observation on terminology. The word "split" is used to designate a pair of fragments of given kind, internal energy, combined spin, and relative angular momentum. The wave functions are discussed in $3A$-dimensional configuration space, where A is the number of nucleons in the system. On occasion the word "channel" is used to designate that part of the $3A$-dimensional space in which the wave functions describing a particular split do not vanish.

2. DEFINITION OF THE COLLISION MATRIX

We work with wave functions which describe a stationary state of A nucleons, with total energy E, and angular momentum indicated by the quantum numbers J(total) and M(z-component). These functions depend on spin variables and the $3A$ position coordinates of the nucleons. Because of the short range of nuclear forces the $3A$-dimensional configuration space may be supposed to be divided into two regions by the closed hypersurface S defined by Wigner and Eisenbud (6). S is such that at all points in its interior the wave function must describe all nucleons in nuclear interaction, while at points exterior to S, the wave function describes the system split into two or more groups of nucleons with no nuclear interaction between the groups.

This discussion is limited to the case that only two subnuclei appear in the external region. However this splitting may be done in many ways, so that the total Ψ will be a sum of functions each of which describes the system in one particular splitting. We write, again following Wigner and Eisenbud,

$$\Psi(JME) = \sum_{s,\,l} F_{sl}(E - \epsilon_s, r_s) \sum_{m} S(j_s, l)_{JM.\,m_s} Y_l^m(\omega_s)\chi_{sm_s}. \tag{2.1}$$

Here F_{sl} is the radial wave function of the relative coordinates of the centers of mass of the two fragments formed by the split s; $Y_l^m(\omega_s)$ is the usual tesseral harmonic function of the relative angular coordinates. $\chi_{s\nu}$ is the product of the internal wave functions of the two fragments, in which s describes the actual groups of nucleons, together with their internal energies which sum to ϵ_s, and their spins which combine to form a total spin of j_s, while $\nu = M-m$ specifies the z-component of this combined spin. The sum over s includes a sum over all possible values of ϵ_s and j_s. The coefficients $S(j_s, l)_{JM;m\nu}$ are given by Wigner (5, Chap. 8, Section 6). They specify the correct combinations of components $\nu = M-m$ of spin, and m of relative angular momentum to give a system with total angular momentum J with z-component M.

The function Ψ is supposed to be properly antisymmetrized in the coordinates of all nucleons. Outside S partial interchange of nucleons between fragments will cause Ψ to vanish, since the internal wave functions of the fragments do not overlap. Interchange of nucleons within fragments is included in the internal functions. If the two fragments be identical, then Ψ must have the proper symmetry with respect to their interchange as units. This will affect the range of possible values of l.

Our discussion is chiefly concerned with the radial wave functions. These are expressed in terms of the usual incoming wave

$$r_s^{-1}\phi_s^*(r_s),$$

and outgoing wave

$$r_s^{-1}\phi_s(r_s).$$

These satisfy the one-dimensional Schroedinger equation

$$\frac{h^2}{2\mu_s}\phi_s'' + (E - \epsilon_s - v_s)\phi_s = 0 \tag{2.2a}$$

where v_s is the external potential for the coulomb and centrifugal forces. The functions ϕ_s are normalized to unit current; i.e.

$$(2i\mu_s)^{-1}h\left(\phi_s^* \frac{d\phi_s}{dr_s} - \phi_s \frac{d\phi_s^*}{dr_s}\right) = 1. \tag{2.2b}$$

In these and following equations h is Planck's constant divided by 2π.

F_{sl} is a linear combination of these incoming and outgoing waves; we will write

$$\Psi(EJM) = \sum_{sl} [a_{sl}\phi_{sl}^* - (\sum_{s'l'} U_{sl;s'l'}^J a_{s'l'})\phi_{sl}]r_s^{-1}\psi_{sl}. \tag{2.3}$$

Here ψ_{sl} denotes the sum over m contained in (2.1); these functions may be taken to be real, and are orthogonal on S or on any similar hypersurface enclosing S. They are to be so normalized that (see (6))

$$\int_S r_s^{-1}r_{s'}^{-1}\psi_{sl}\psi_{s'l'}dS = \delta_{ss'}\delta_{ll'}. \tag{2.4}$$

This integral also includes integration over internal coordinates.

The coefficients a_{sl} are arbitrary; they define the amplitudes of the incoming waves. The outgoing amplitudes are then determined by the collision matrix $\mathbf{U}^J$. The existence and uniqueness of $\mathbf{U}^J$ follows from the assumption that H is separable in the asymptotic region. For if H is separable there exists a unique Ψ defined by the boundary condition that there be unit incoming current in the (sl)-split and no other incoming current. Continuation through the internal region will, in principle, uniquely determine the outgoing waves in this particular Ψ; i.e. will uniquely determine the column $U^J_{sl,s'l'}$ of $\mathbf{U}^J$. Equation (2.3) is then a linear combination of these special functions. $\mathbf{U}^J(E)$ depends on the total angular momentum J, and the total energy E of the system. In the absence of external fields it is supposed to be independent of M.

In the following discussion attention is confined to a single system, so the indices E, J, M will be omitted. Also the pair (s,l) of indices will be indicated by one index, k. Thus we consider the wave function which, in the asymptotic region, has the form

$$\Psi = \sum_k [a_k \phi_k^* - (\sum_m U_{km} a_m)\phi_k] r_k^{-1} \psi_k. \tag{2.3a}$$

It is well known that the matrix $\mathbf{U}$ is both unitary and symmetric, so the proof is merely sketched here. Since both Ψ of (2.3a) and its complex conjugate must satisfy the Schroedinger equation for a steady state it follows from the definition of $\mathbf{U}$ that

$$\sum_{jm} U_{kj}^* U_{jm} a_m = a_k. \tag{2.5}$$

The system is stationary, so the total current, summed over all splits, must vanish.

$$\sum_k [|a_k|^2 - |\sum_m U_{km} a_m|^2] = 0. \tag{2.6}$$

Since the a_k are arbitrary complex numbers it quickly follows from (2.5) and (2.6) that

$$\sum_j U_{kj}^* U_{jm} = \delta_{km} \text{ and } U_{kj} = U_{jk} \tag{2.7}$$

so that $\mathbf{U}$ is both unitary and symmetric. Then so also is the matrix

$$\mathbf{V} = \Omega^{-1}\mathbf{U}\Omega^{-1} \tag{2.8a}$$

if

$$\Omega_{jk} = \delta_{jk} \exp i\omega_k; \tag{2.8b}$$

i.e. if Ω is any unitary diagonal matrix.

Then there exists a complete set of real orthonormal eigenvectors†, $g^{(k)}$, of $\mathbf{V}$, belonging to the eigenvalues $\exp(2i\tau_k)$. We can then write

†*Since* $\mathbf{V}$ *is unitary* $\mathbf{V}^{-1} = \mathbf{V}\dagger$, *but* $\mathbf{V}$ *is also symmetric, so* $\mathbf{V}^{-1} = \mathbf{V}^*$ (*the complex conjugate of* $\mathbf{V}$). *Then* $\mathbf{V}g = e^{2i\tau}g$ *implies* $\mathbf{V}^*\mathbf{V}g = g = e^{2i\tau}\mathbf{V}^*g$. *The complex conjugate of this last equation is* $\mathbf{V}g^* = e^{2i\tau}g^*$. *Hence either* g *is a real vector, or* g *and* g^* *are degenerate eigenvectors of* $\mathbf{V}$ *which may be replaced by a pair of real orthonormal eigenvectors belonging to the value* $e^{2i\tau}$.

$$\mathbf{U} = \Omega\left[\sum_{k=1}^{N} \exp(2i\tau_k)\mathbf{g}^{(k)}\times\mathbf{g}^{(k)}\right]\Omega \tag{2.9}$$

$$= \Omega\left[\mathbf{1} + 2i\sum_{k=1}^{N} (\cot\tau_k - i)^{-1}\mathbf{g}^{(k)}\times\mathbf{g}^{(k)}\right]\Omega.$$

The matrix element

$$(g^{(k)}\times g^{(k)})_{jm} = g_j^{(k)} g_m^{(k)};$$

N is the number of possible splittings, i.e. the dimension of $\mathbf{U}$. Not all τ_k need be distinct.

The form (2.9) is useful for comparison with the form (3.16) below, obtained by application of the compound nucleus hypothesis.

3. THE "COMPOUND NUCLEUS" ASSUMPTION

The general form for the collision matrix has so far been expressed only in such a complicated way (6) that one naturally seeks some simplifying assumption. In the general work of Wigner the concept of the compound nucleus is used only in the weak sense that all nucleons enter the interior region. This assumption is generally understood in a much stronger sense which takes account of the strength of nuclear forces. For the purpose of the present discussion we use the following strong assumption.

If there is no ingoing current in the split k the outgoing current in that split depends only on the rate of formation of the compound nucleus and is independent of the way that nucleus is formed. (*A*)

Moreover the density of matter in the interior of the compound nucleus, i.e. the value of $|\Psi|^2$ in the interior of the hypersurface S, is also independent of the way the nucleus is formed. (*B*)

This seems to be a fair statement of the concept of the compound nucleus as it is used in the literature. It is not maintained that this is a description which is always applicable, but there is some evidence, both theoretical and experimental, that it is a good approximation in many cases.

The usual justification is based on statistical concepts (2); that the life of the compound nucleus is so long on a nuclear time scale (in comparison with the time for a nucleon of average internal kinetic energy to traverse the nucleus) that statistical equilibrium is attained. Thus the internal density depends only on E, J, M, the constants of motion of the system. So also the subsequent mode of disintegration will be independent of the mode of formation.

The general theory of the collision matrix shows that these statements are approximately true in the neighborhood of a resonance. However this theory contains no assumptions about the strength of nuclear forces; it is sufficiently general to cover all cases. Hence that it does not confirm the compound nucleus concept is not a refutation of the statistical arguments which are valid only for strong nuclear forces.

Despite the frequent use of the compound nucleus idea there is very little clear-cut experimental evidence defining its region of validity. Some examples of available evidence will be discussed in the next section.

We now seek the form of the collision matrix determined by assumptions A and B. As a preliminary we define a "normal system". This shall be one for which the compound nucleus is formed at unit rate; i.e. there is unit flow of matter current into the region interior to the surface S. If, in a normal system, there is no ingoing current in the k-split, the outgoing current in that split will be designated by G_k^2. Assumption A states that G_k^2 is a function of energy which is independent of the choice of those splits other than k which contain the ingoing currents. On the other hand, if the j-split contains the only ingoing current, the compound nucleus is formed at unit rate (i.e. the system is normal) if

$$|a_j|^2(1 - |U_{jj}|^2) = 1 - G_j^2. \tag{3.1}$$

This follows since the net influx (cf. equation (2.3a)) is $|a_j|^2(1 - |U_{jj}|^2)$ and this must be less than unity by the amount G_j^2 of outgoing current which would arise if the current were ingoing in some other split. In using (3.1) it will be assumed that $|U_{jj}|^2 < 1$, i.e. that absorption can occur in any energetically possible split.

In the normal system for which $|a_j|^2$ is given by (3.1) and $a_k = 0$, $k \neq j$, the outgoing current in the k-split is $|U_{kj}|^2 |a_j|^2$ and this equals G_k^2 by definition. Hence if $j \neq k$,

$$|U_{kj}|^2 = G_k^2(1 - G_j^2)^{-1}(1 - |U_{jj}|^2). \tag{3.2}$$

From the unitary properties of $\mathbf{U}$,

$$\sum_k |U_{kj}|^2 = 1 \text{ and } \sum_j |U_{kj}|^2 = 1,$$

we find from (3.2) that

$$\sum_k G_k^2 = 1, \tag{3.3}$$

$$|U_{kj}|^2 = BG_k^2G_j^2, \quad k \neq j, \tag{3.4a}$$

$$|U_{kk}|^2 = 1 - BG_k^2(1 - G_k^2); \tag{3.4b}$$

where

$$B = \sum_k (1 - G_k^2)^{-1}(1 - |U_{kk}|^2). \tag{3.4c}$$

These equations give the moduli $|U_{kj}|$ of the elements of $\mathbf{U}$ in terms of the energy functions G_k^2 and B, all of which are real and positive. We now put $U_{kj} = |U_{kj}| \exp i\omega_{kj}$, and seek the values of the ω_{kj}. These can be partly determined by a discussion of a normal system in which ingoing current flows in two channels.

If $a_j \neq 0$, $a_k \neq 0$, while all other a's are zero, the system is normal if

$$|a_k|^2 - |U_{kk}a_k + U_{kj}a_j|^2 + |a_j|^2 - |U_{jj}a_j + U_{jk}a_k|^2 = 1 - G_j^2 - G_k^2 \tag{3.5}$$

and for $i \neq j$, $i \neq k$,

$$|U_{ik}a_k + U_{ij}a_j|^2 = G_i^2. \tag{3.6}$$

For convenience we put

$$a_k = (\sqrt{B}G_k)^{-1}e^{i\nu_k}c_k,$$

c_k real, and write a_j in a similar fashion. Using the values (3.4) for $|U_{kj}|$, Equation (3.5) yields a value for the quantity

$$(2c_kc_j)^{-1}(1 - c_k^2 - c_j^2),$$

which, when substituted into (3.6), gives

$$\begin{aligned}(1 - G_k^2 - G_j^2)\cos(\omega_{ik} - \omega_{ij} + \nu_k - \nu_j)\\ + (B^{-1} - G_k^2 + G_k^4)^{-\frac{1}{2}}\cos(\omega_{kk} - \omega_{jk} + \nu_k - \nu_j)\\ + (B^{-1} - G_j^2 + G_j^4)^{-\frac{1}{2}}\cos(\omega_{jj} - \omega_{jk} + \nu_j - \nu_k) = 0.\end{aligned} \tag{3.7}$$

The first observation is that $\omega_{ik} - \omega_{ij} = \pm\lambda_{kj} \pm 2n\pi$; in which λ_{kj} is some quantity which does not depend on the index i. We are interested only in $\exp i(\omega_{ik} - \omega_{ij})$ so we can restrict ourselves to principal values of this difference. The ambiguity of sign, which in (3.7) may depend on i, is removed if we consider normal systems with more than two incoming currents. So we may write

$$\omega_{ik} - \omega_{ij} = \lambda_{kj}. \tag{3.8}$$

We may take λ_{jk} as given and consider (3.8) as a set of linear equations for the ω_{ik}. Let σ_1 be an arbitrary real number, and define $\sigma_k = \sigma_1 + \lambda_{k1}$. It then follows that $\lambda_{jk} = \sigma_j - \sigma_k$. The determinant of equations (3.8) vanishes, so we can choose at least one ω_{ik} arbitrarily. Put $\omega_{12} = \sigma_1 + \sigma_2$. A simple induction then establishes that

$$\omega_{ik} = \sigma_i + \sigma_k, \qquad i \neq k. \tag{3.9}$$

Now we can find the values σ_j in terms of ω_{jj} and other quantities previously defined. For if we choose

$$a_k = (2\sqrt{B}G_k)^{-1}e^{-i\sigma_k},\ a_j = (2\sqrt{B}G_j)^{-1}e^{-i\sigma_j},$$

it is clear that (3.6) is satisfied for every i distinct from j and k; i.e. by definition these values of a_k, a_j define a normal system so that (3.5) must also be satisfied. This requires, for this a_k, a_j, and using (3.9),

$$\begin{aligned}[B^{-\frac{1}{2}}|U_{kk}|\cos(\omega_{kk} - 2\sigma_k) + \tfrac{1}{2} - G_k^2]\\ = -[B^{-\frac{1}{2}}|U_{jj}|\cos(\omega_{jj} - 2\sigma_j) + \tfrac{1}{2} - G_j^2].\end{aligned} \tag{3.10}$$

Equation (3.10) must hold for every choice of j, k; consequently each side must vanish.

If we define

$$\cos\rho_k = -|U_{kk}|^{-1}(\tfrac{1}{2} - G_k^2)\sqrt{B} \tag{3.11a}$$

$$\sin\rho_k = |U_{kk}|^{-1}\sqrt{1 - B/4} \tag{3.11b}$$

then we find $\exp i(\omega_{kk} - 2\sigma_k) = \exp \pm i\rho_k$. The sign before ρ_k remains ambiguous because (3.10) defines only the real part of this exponential. A priori

one might suppose that the more general equation (3.7) would be satisfied only for special values of the phase difference $\nu_k - \nu_j$. However it will be observed that a solution of the form above exists, for any value of $\nu_k - \nu_j$, provided that one chooses the same sign for every ρ_k. We then write

$$\exp i(\omega_{kk} - 2\sigma_k) = \exp i\rho_k. \tag{3.12}$$

To change sign of all ρ_k is merely a matter of notation.

By definition ω_{jk} are real, and so ρ_k must also be real. Equation (3.11*b*) then provides an upper bound for the positive function B;

$$0 \leqslant B \leqslant 4. \tag{3.13}$$

This bound will be made explicit if we define a new function $\tau(E)$ by the relations

$$\sin \tau = \sqrt{B/4}, \quad \cos \tau = \sqrt{1 - B/4}. \tag{3.14}$$

To summarize, using the values (3.4) for $|U_{jk}|$; the form (3.9), with σ_k given by (3.12), for the arguments of off diagonal elements; and the definition (3.14),

$$U_{kj} = 2 \sin \tau G_k G_j \exp \tfrac{1}{2}i(\omega_{kk} + \rho_k + \omega_{jj} + \rho_j), \qquad k \neq j, \tag{3.15a}$$

$$U_{kk} = [1 - 4 \sin^2 \tau (G_k^2 - G_k^4)]^{\frac{1}{2}} \exp i\omega_{kk}. \tag{3.15b}$$

These results can be expressed in the matrix relation

$$\mathbf{U} = \Omega[\mathbf{1} + 2i(\cot \tau - i)^{-1}\mathbf{G} \times \mathbf{G}]\Omega \tag{3.16a}$$

where Ω is diagonal with elements

$$\Omega_k = \exp \tfrac{1}{2}i(\omega_{kk} + \rho_k - \tau - \tfrac{1}{2}\pi), \tag{3.16b}$$

and the matrix $\mathbf{G} \times \mathbf{G}$ has elements

$$(G \times G)_{kj} = G_k G_j. \tag{3.16c}$$

The form (3.16*a*) is to be compared with the general form (2.9); we see that part A of the compound nucleus assumption implies that there is a decomposition of $\mathbf{U}$ of the form $\Omega\mathbf{V}\Omega$ such that $\mathbf{V}$ has just two eigenvalues; the value 1, which is $(N - 1)$ fold degenerate, and the value $e^{2i\tau}$, for which the eigenvector has the components G_k. This also gives the meaning of the function $\tau(E)$ defined by equations (3.14).

From part B of the assumption we obtain more specific values for Ω_k and G_k. According to B the value of $|\Psi|^2$ just inside the surface S is independent of the choice of incoming currents. It may be observed in passing that this assumption about surface values ensures that small changes in the definition of the surface S (i.e. in the definition of nuclear radii) do not affect the statement that

$$\int_V |\Psi|^2 dv$$

is independent of the choice of incoming current. (V is the volume interior to S.) Moreover the physical implications of the assumption are reasonable; it states that the probability of finding particles of type k nascent on the nuclear surface is the same, whether these particles initiate or result from the reaction. More

specifically, if the two reactions $Z^{A-1}(n,p)\ (Z-1)^{A-1}$ and $(Z-1)^{A-1}(p,n)Z^{A-1}$ go through a compound nucleus with the same energy E and angular momentum J, then the chance of finding a neutron of given angular momentum on the surface of the compound nucleus is the same in each case.

To use assumption B, we note that Ψ is continuous at the nuclear surface and so may be evaluated there in energetically possible channels either from the form (2.3a) or from internal values. We consider two normal systems, first with incoming current only in the k-split, second only in the j-split. The appropriate value of $|a_k|^2$ in the first case is (cf. 3.1)

$$|a_k|^2 = (2G_k \sin \tau)^{-2}$$

with a similar expression for $|a_j|^2$. Since $|\Psi|^2$ is to be the same in the k-channel in each case, we require at S,

$$G_k^{-2}|\phi_k^* - U_{kk}\phi_k|^2 = G_j^{-2}|U_{kj}\phi_j|^2. \tag{3.17}$$

For convenience we write the outgoing radial wave function of the k-split, evaluated at S where $r_k = a_k$

$$|\phi_k(a_k)| = f_k;\ \ \phi_k(a_k) = f_k \exp(i\eta_k) \tag{3.18a}$$

and its derivative

$$\phi_k'(a_k) = (q_k + ip_k)\phi_k(a_k). \tag{3.18b}$$

Then with U_{kk}, U_{kj} given by (3.15), we find from (3.17)

$$\cos(\omega_{kk} + 2\eta_k) = \cos(\tau - \rho_k + \tfrac{1}{2}\pi).$$

Again this is insufficient to determine the relative signs of the angles, but consideration of a normal system with incoming currents in two splits yields

$$\Omega_k = \exp \tfrac{1}{2}i(\omega_{kk} + \rho_k - \tau - \tfrac{1}{2}\pi) = \exp(-i\eta_k). \tag{3.19}$$

For the next step of the argument we consider two functions, Ψ_1 and Ψ_2, both having the form (2.3a) in the external region, differing only in having different energies E_1 and E_2. We suppose that no new reactions become possible between E_1 and E_2, i.e. that the dimension of $\mathbf{U}$ remains constant over this interval.

Let V denote the volume interior to S. Form

$$\begin{aligned}(E_2 - E_1)\int_V \Psi_1^\dagger \Psi_2 dv &= \int_V (\Psi_1^\dagger H\Psi_2 - \Psi_2 H\Psi_1^\dagger)dv \\ &= \sum_k \frac{h^2}{2\mu_k}(F_{2k}F_{1k}^{*\prime} - F_{1k}^* F_{2k}').\end{aligned} \tag{3.20}$$

This† results from the application of Green's theorem in $3A$ dimensions; the

†*Properly speaking, the normal derivative on S is not constant over that part of S which corresponds to the k-split. The value on the right of* (3.20) *for*

$$\int_S \frac{h^2}{2\mu}\left(\Psi_2 \frac{\partial \Psi_1^*}{\partial n} - \Psi_1^* \frac{\partial \Psi_2}{\partial n}\right)dS$$

is to be taken as an application of the mean value theorem.

volume integral is transformed into an integral over S (cf. 6). The sum on the left runs over all splits, including the energetically impossible ones. F_k represents the radial wave function of the k-split, evaluated on S.

If now we put $(E_2 - E_1) = \Delta E$ and form a Taylor expansion about the energy $E = E_1$, the coefficients of each power of ΔE in (3.20) must separately vanish. The terms independent of ΔE state only that the net current is zero. The terms in the first power of ΔE give

$$\int_V |\Psi|^2 dv = \sum_k' \frac{h^2}{2\mu_k}\left(F_k^{*\prime}\frac{dF_k}{dE} - F_k^*\frac{dF_k'}{dE}\right) - T(E). \tag{3.21}$$

Here the summation extends only over the energetically possible splits. The similar sum over impossible splits is represented by $T(E)$. We will not here attempt to evaluate $T(E)$; it represents the integral of $|\Psi|^2$ over impossible channels in the region outside S. Hence it is properly combined with $\int_V |\Psi|^2\, dv$ to give the probability that the nucleons are all "close together". Normally $T(E)$ will be a small quantity, since Ψ will decay approximately exponentially in impossible channels outside S. One might expect $T(E)$ to become large at certain discrete energies, but this question is not part of the present problem.

We define a normal system with incoming current only in the k-split, and for convenience take

$$a_k = (2iG_k \sin \tau)^{-1} \exp - i(\tau - \eta_k).$$

Then from (3.15) and (3.19) we find in possible channels on S

$$\begin{aligned} &F_j = - U_{kj}a_k\phi_j(a_j) = - G_jf_j, \quad F_j' = - (q_j + ip_j)G_jf_j,\ j \neq k; \\ &F_k = a_k(\phi_k^* - U_{kk}\phi_k) = - G_kf_k, \\ &F_k' = - (q_k + ip_k)G_kf_k - G_k^{-1}(\cot \tau - i)f_kp_k. \end{aligned} \tag{3.22}$$

For this normal system, Equation (3.21) becomes

$$\left[\int_V |\Psi|^2 dv + T(E)\right]_N = - \tfrac{1}{2}h\left[\sum_j G_j^2 p_j^{-1}\frac{dq_j}{dE} + \frac{d}{dE}\cot \tau - \frac{d}{dE}\log p_k^{-1}G_k^2\right]. \tag{3.23}$$

In simplifying (3.23) we have used the relation

$$\frac{h}{\mu_j} f_j^2 p_j = 1 \tag{3.24}$$

which results from the normalization (2.2) of the ϕ_j.

The subscript N on the left side of (3.23) indicates that this relation holds for a normal system; by assumption B this quantity is independent of k, the split containing incoming current. On the right side there remain terms which appear to depend on k. To satisfy B we must therefore require

$$\frac{d}{dE}\log p_k^{-1}G_k^2 = X(E)$$

where $X(E)$ is a function of energy which does not depend on k. Then

$$p_k^{-1}G_k^2 = \beta_k^2 \exp \int X(E)dE.$$

Since

$$\sum_j G_j^2 = 1,$$

this result may be written

$$G_k^2 = \beta_k^2 p_k (\sum_j \beta_j^2 p_j)^{-1} \tag{3.25}$$

in which β_k^2 is a positive constant, independent of energy; p_k, the imaginary part of the logarithmic derivative of the outgoing wave at the nuclear surface, appears as a penetration factor.

This form for G_k^2, particularly the constancy of β_k^2, is actually here established only in each energy interval in which the dimension of **U** is constant; i.e. in each interval between thresholds. If we join these solutions by the requirement that G_k^2 be a continuous function of energy, we must require β_k^2 constant for all energies above the k-threshold. This seems to be a natural requirement, since it is unreasonable to assume a discontinuous change in the cross section for, say, proton production at the threshold for production of α-particles. Because $p_j = 0$ at the j-threshold, for every j, the normalization factor

$$(\sum_j \beta_j^2 p_j)^{-1}$$

is continuous, though its derivative is discontinuous at each threshold.

It is now useful to introduce width-functions Γ_k defined by

$$h^{-1}\Gamma_k \left[\int_V |\Psi|^2 dv + T(E) \right] = \left| \sum_j U_{kj} a_j \right|^2. \tag{3.26}$$

This definition states that, if there be no incoming k-current, the outgoing k-current is Γ_k/h times the probability that all particles are in the internal region, or outside but bound in nonsplitting states. When (3.26) is applied to a normal system with $a_k = 0$, we see that the right side becomes G_k^2. Since

$$\sum_k G_k^2 = 1,$$

we then have

$$\left[\int_V |\Psi|^2 dv + T(E) \right]_N = h(\sum_k \Gamma_k)^{-1} = h\Gamma^{-1}, \tag{3.27a}$$

$$G_k^2 = \Gamma_k \Gamma^{-1}. \tag{3.27b}$$

Using the result (3.25) and the definitions (3.27) the differential equation (3.23) becomes

$$\frac{d}{dE} \sum_j \beta_j^2 (q_j + p_j \cot \tau) = -2\Gamma^{-1} \sum_j \beta_j^2 p_j. \tag{3.28}$$

Recalling that cross sections $\sigma_{jk} \sim |\delta_{jk} - U_{jk}|^2$, one sees from (3.15) that thei. maxima occur when $\sin \tau(E_r) = 1$. Thus at a resonance we have†

$$\cot \tau(E_r) = 0. \tag{3.29}$$

Then (3.28) gives us the first nonzero term in the Taylor expansion of $\cot \tau(E)$ near $E = E_r$,

$$\cot \tau(E) = -2\Gamma_r^{-1}(E - E_r) + \ldots. \tag{3.30}$$

In (3.30) we have neglected the change in $\sum_j \beta_j^2 p_j$ and $\sum_j \beta_j^2 q_j$ over the interval $E - E_r$.

Before going on to a discussion of these results it should be remarked that the developments of this section apply when there are more than two energetically possible modes of splitting. It may easily be shown that **U** may be written in the same forms for the case of two possible splittings. When only elastic scattering is possible the statements become trivial.

4. DISCUSSION

It is hardly necessary to point out that the collision matrix derived from the compound nucleus postulate gives cross sections with standard properties in the various energy regions. The connection between the matrices $\mathbf{U}^J$ and experimentally measurable cross sections is given by Wigner and Eisenbud (6). The differential cross sections take a somewhat complicated form because of the necessity of transforming back from wave functions of specified total J, M to those of specified ν,m. However the total cross section takes the simple form

$$\sigma_{ss'} = \frac{4\pi}{k_s^2(2j_s + 1)} \sum_{ll'J} (2J + 1)|\delta_{ss'}\delta_{ll'} - U_{sl,s'l'}|^2$$

in which we have reverted to the notation (sl) in place of the single index k used to denote a split in Section 3; k_s is the wave number of the incident particles; the rest of the notation is explained in Section 2.

Using the results of the last section (Equations 3.16 and 3.19) this becomes

$$\sigma_{ss'} = \frac{4\pi}{k_s^2(2j_s + 1)} \sum_{ll'J} (2J + 1)\left|\delta_{ss'}\delta_{ll'}e^{i\eta_{sl}} \sin \eta_{sl} + \frac{G_{sl}^J G_{s'l'}^J}{\cot \tau^J - i}\right|^2. \tag{4.1}$$

In (4.1) the sum over J extends over all the compound nuclei which can be formed by different combinations of angular momenta and spin of the target-particle system s. For each different compound nucleus (different J-value) there will be a vector G_{sl}^J and a characteristic nuclear phase-shift τ^J.

The potential scattering term $e^{i\eta_{sl}} \sin \eta_{sl}$ in (4.1) is standard; it represents the scattering by a hard sphere. The penetration factor p_{sl} contained in

$$G_{sl}^2 = \Gamma^{-1}\Gamma_{sl} = p_{sl}\beta_{sl}^2\Big(\sum_{s'l'} \beta_{s'l'}^2 p_{s'l'}\Big)^{-1}$$

is also well known. For example, if one calculates a simple spherical barrier problem, the transmission factor is just $p|\Psi|^2$, in which Ψ is the internal wave

†*Compare with the appendix; this definition of resonance is also obtained by specializing Wigner's general collision matrix.*

function evaluated at the barrier. This penetration factor gives the usual behavior for absorption and scattering of slow particles (including the v^{-1} rule for s-neutrons) and for emission near threshold.

For cross sections near resonance we use the differential equation (3.28), and in particular the resultant Taylor expansion (3.30). Assuming that resonance occurs in only one of the possible compound nuclei (one J-value) at a given energy E_r, we find from (4.1) and (3.30)

$$\sigma_{ss'} \approx \frac{4\pi}{k_s^2}\frac{(2J+1)}{(2j_s+1)}\sum_{ll'}\left|\delta_{ss'}\delta_{ll'}e^{i\eta_{sl}}\sin\eta_{sl} + \frac{(\Gamma_{sl}^J\Gamma_{s'l'}^J)^{\frac{1}{2}}}{2(E-E_r^J+\frac{1}{2}i\Gamma^J)}\right|^2 \tag{4.2}$$

i.e., a standard Breit–Wigner one-level form.

At higher energies cross sections tend to vary more slowly. This is usually stated as "the energy levels become dense and wide". In our description, this is equivalent to saying that $\Gamma(E)$ becomes large and varies slowly at high energies. Since the p_{sl} vary like k_s, and q_{sl} tend to zero at high energies, the differential equation (3.28) shows that cot τ has no poles or zeros at such high energies. In this region one can apply (4.1) for cross sections, with the assumption that the cot τ^J are slowly varying monotonic functions.

Feshbach, Peaslee, and Weisskopf (3), (called *FPW* in the following few paragraphs) have calculated cross sections for scattering and absorption, using the compound nucleus hypothesis and a phenomenological approach. It is clear that our results should be similar.

Their discussion centers around the logarithmic derivative of the radial wave function for the split which contains the incoming current. If we let that be the sl-split, the *FPW* function $f(E)$ becomes in our notation

$$f(E) = f_0 - ih = a_s[q_{sl} + p_{sl}G_{sl}^{-2}\cot\tau - ip_{sl}G_{sl}^{-2}(1-G_{sl}^2)]$$

where a_s is the value of r_s at S.

The further identifications are obvious (*FPW* notation on left)

$$k_s\left|\frac{v'}{v}\right|\cos(\delta'-\delta) = q_{sl}$$

$$k_s\left|\frac{v'}{v}\right|\sin(\delta'-\delta) = p_{sl}.$$

Then the condition of resonance (*FPW* Equation 40) becomes

$$q_{sl} + p_{sl}G_{sl}^{-2}\cot\tau(E_r) = q_{sl}$$

i.e. cot $\tau(E_r) = 0$

which is also our condition of resonance.

FPW widths are defined in terms of df_0/dE, which can be calculated using (3.25) for G_{sl}^2, and the differential equation (3.28), to be

$$\frac{df_0}{dE} = \frac{a_s}{\beta_{sl}^2}\left[-\frac{2}{\Gamma}\sum_{s'l'}\beta_{s'l'}^2 p_{s'l'} - \sum_{s'l'\neq sl}\beta_{s'l'}^2 q_{s'l'}\right].$$

If one neglects the sum over $\beta^2_{s'l'}q_{s'l'}$, which is small, the width definitions *FPW*(41) and *FPW*(12) become

$$\Gamma_{sl} = G^2_{sl}\Gamma$$

and

$$\Gamma_{abs} = (1 - G^2_{sl})\Gamma$$

which are also our definitions. To obtain these results it is necessary to remember that our outgoing wave function ϕ is normalized differently from that of *FPW*.

These width functions are defined only at resonance by *FPW*, but can be extended to agree with ours at all energies and are thus related to the integral of $|\Psi|^2$ over the nuclear interior (cf. Equation (3.27)).

Finally it may be noted that the function $\tan z(W)$ defined by *FPW*(19) for neutrons of zero angular momentum is related to $\cot \tau$ in the obvious way.

$$-K \tan z(W) = G^2_{n0}k_n \cot \tau$$

where we have used $q_{n0} = 0$, $p_{n0} = k_n$ for such neutrons. This relation is the one to be expected between the phase shift τ of the external wave function caused by nuclear effects, and the phase $(z(W) - \frac{1}{2}\pi)$ of the internal wave function evaluated at the nuclear surface.

The value of the compound nucleus hypothesis lies in the considerable simplification it brings to the formulae for cross sections. The whole effect of the nuclear interaction is expressed in the one function $\tau(E)$ and the constants β_{sl}.

It thus seems well worth while to inquire of experiment when this simplifying assumption can be used. Most physicists would agree that it applies in the region of a resonance, but this is insufficient for the calculations made here. To make them valid the hypothesis must hold over the whole range of energy in which one wishes, for example, to treat the β_{sl} as constants.

There are some results which suggest that this is possible. For example, Ghoshal (4) has found that the reactions $Ni^{60} + a$, $Cu^{63} + p$, over a considerable range of energies, produce n, $2n$, pn in equal ratios. This suggests that the hypothesis holds for the nucleus Zn^{64} even at such energies that threefold splittings are possible. Beiduck, Pruett, and Konopinski (1) have shown that the complex energy dependence of the cross sections for $D(dp)T$ and $D(dn)He^3$ can be explained by supposing that a fixed set of compound nuclei (set of J-values) is formed. This is a use of the compound nucleus hypothesis for a system of only four particles, where the hypothesis might also be doubted.

The usual criterion for the formation of a compound nucleus is just our assumption A; that the branching ratio of outgoing currents is independent of the incoming currents. The present calculations give another more stringent result. It becomes more evident if we write the reaction cross section (4.1) for $s' \neq s$

$$\sigma_{ss'} = \frac{4\pi}{k_s^2(2j_s + 1)} \sum_{ll'J} (2J + 1)(G^J_{sl})^2(G^J_{s'l'})^2 \sin^2 \tau^J. \tag{4.3}$$

At a resonance $\sin^2\tau^J(E_r) = 1$; $\cot \tau^J(E_r) = 0$. However the differential equation (3.28) shows that the slope of

$$\sum_{sl}\beta_{sl}^2(q_{sl} + p_{sl}\cot\tau^J)$$

is always negative. The q_{sl} and p_{sl} are finite, slowly varying functions of E. Hence we deduce that $\tau^J(E)$ is essentially a monotomic function of E and resonances occur at $\tau^J(E) = (2n + 1)\pi/2$. This requires that τ^J pass through a value $n\pi$ between each pair of resonances, and so every reaction cross section (but not scattering) vanishes at this intermediate energy.

If one could find a reaction in which only one J-value could be formed over a range of energy which includes two resonances, all reaction cross sections should vanish at some energy between the resonances. If an experiment could be devised to measure the reaction or absorption cross sections of one particular compound nucleus, the presence (or absence) of these zeros would be striking evidence for (or against) the value of the compound nucleus hypothesis.

The writer wishes to thank the members of the staff of Theoretical Physics branch at this laboratory for their continued patience through many conversations on this subject, and Professors E. P. Wigner, W. H. Watson, and J. D. Jackson who read and helpfully commented on a preliminary draft of this discussion.

APPENDIX

Here we will show that the compound nucleus hypothesis requires a certain special case of Wigner's collision matrix. The discussion in Section 4 shows that this special case also corresponds to the work of Feshbach, Peaslee, and Weisskopf.

The basic step in determining the general form of the collision matrix is to evaluate the "derivative matrix" $\mathbf{R}$ which relates (cf.(7)) the values of the radial wave functions of the various splits to the values of their derivatives at the surface S; in the notation of Section 3,

$$\frac{F_k}{\sqrt{2\mu_k}} = \sum_j \frac{R_{kj}F_j'}{\sqrt{2\mu_j}} \tag{A.1}$$

and it is shown by Wigner and Eisenbud (6) that

$$R_{kj} = \sum_\lambda \frac{\gamma_{\lambda k}\gamma_{\lambda j}}{E_\lambda - E}. \tag{A.2}$$

The $\gamma_{\lambda k}^2$ are energy independent reduced widths of the nuclear level of energy E_λ.

Given $\mathbf{R}$, the matrix $\mathbf{U}$ can be found in various ways of which the following is perhaps most interesting for the special case examined here. Write $\mathbf{U} = \Omega\mathbf{V}\Omega$, where the elements

$$\Omega_k = \exp(i\omega_k)$$

of the diagonal matrix Ω are to be chosen later. Let g_k represent the components of a real eigenvector of $\mathbf{V}$, belonging to the eigenvalue $e^{2i\tau}$. Then consider a state with incoming currents in every split given by

$$(A.3) \qquad a_k = \tfrac{1}{2}\exp(-i\omega_k)\, e^{-i\tau} g_k$$

so that

$$(A.4) \qquad (\mathrm{U}a)_k = \tfrac{1}{2}i\exp(i\omega_k)\, e^{i\tau} g_k.$$

Then with the notation

$$\phi_k = f_k \exp(i\eta_k), \qquad \phi_k' = (q_k + ip_k)\phi_k$$

for the outgoing radial wave function and its derivative at S in the k-channel, the radial wave function and its derivative become

$$(A.5) \qquad \begin{aligned} F_k &= g_k f_k \sin(\tau + \omega_k + \eta_k) \\ F_k' &= F_k[q_k + p_k \cot(\tau + \omega_k + \eta_k)]. \end{aligned}$$

The procedure is to insert $(A.5)$ into $(A.1)$ and identify the Cayley transform† of $\mathbf{V}$ by the fact that the vector g is also an eigenvector of the transform, which belongs to the eigenvalue $(-\cot\tau)$. Inspection of $(A.5)$ shows that this eigenvalue will occur if we choose

$$(A.6) \qquad \omega_k = -\eta_k.$$

When $(A.5)$ is substituted into $(A.1)$, using $(A.6)$ and the relation

$$\mu_k^{-1} h f_k^2 p_k = 1,$$

we find

$$(A.7) \qquad \mathbf{p}^{-\frac{1}{2}}(\mathbf{q} - \mathbf{R}^{-1})\mathbf{p}^{-\frac{1}{2}} g = -\cot\tau\, g$$

where $\mathbf{p}^{-\frac{1}{2}}$ and $\mathbf{q}$ are diagonal matrices with elements $p_k^{-\frac{1}{2}}$ and q_k respectively. Since $(A.7)$ holds for every eigenvector g of $\mathbf{V}$, and the corresponding $\cot\tau$, we can identify the matrix on the left as the Cayley transform of $\mathbf{V}$, and so obtain

$$(A.8) \qquad \mathbf{U} = \Omega[\mathbf{1} - \mathbf{p}^{\frac{1}{2}}\mathbf{R}\mathbf{p}^{\frac{1}{2}}(i\mathbf{1} + \mathbf{p}^{-1}\mathbf{q})]^{-1}[\mathbf{1} + \mathbf{p}^{\frac{1}{2}}\mathbf{R}\mathbf{p}^{\frac{1}{2}}(i\mathbf{1} - \mathbf{p}^{-1}\mathbf{q})]\Omega.$$

This matrix is expressible in more symmetrical form, but $(A.8)$ suffices to show its general nature.

We now wish to find the special case of $(A.8)$ which corresponds to the compound nucleus hypothesis. The matrix $\mathbf{V}^c$ of Equation (3.16) (the superscript c is used here to indicate the special case) has an $(N-1)$ fold degenerate eigenvalue 1, so we cannot form its Cayley transform. However we can form the inverse of this transform, which is

$$i(\mathbf{V}^c + \mathbf{1})^{-1}(\mathbf{V}^c - \mathbf{1}) = -\tan\tau\, \mathbf{G}\times\mathbf{G}.$$

We require that the inverse of the matrix on the left of $(A.7)$ shall have this value; i.e.

$$\mathbf{p}^{\frac{1}{2}}(\mathbf{q} - (\mathbf{R}^c)^{-1})^{-1}\mathbf{p}^{\frac{1}{2}} = -\tan\tau\, \mathbf{G}\times\mathbf{G}.$$

Recalling the value (3.25) for G_k^2, this may be manipulated into the form

†*The Cayley transform* $\mathbf{C}$ *of a unitary matrix* $\mathbf{V}$ *is hermitian. It is given by* $\mathbf{C} = -i(\mathbf{V} - \mathbf{1})^{-1} \times (\mathbf{V} + \mathbf{1})$. *Conversely, if* $\mathbf{C}$ *is given,* $\mathbf{V} = (\mathbf{C} + i\mathbf{1})^{-1}(\mathbf{C} - i\mathbf{1})$. *A unitary matrix* $\mathbf{V}$ *can be formed in this way from any hermitian* $\mathbf{C}$, *but* $\mathbf{C}$ *can be formed from a given* $\mathbf{V}$ *only if the latter has no eigenvalue* 1.

$$(A.9) \qquad R^c_{kj} = \sum_\lambda \frac{\gamma^c_{\lambda k}\gamma^c_{\lambda j}}{E_\lambda - E} = \frac{\beta_k\beta_j}{\sum_l \beta^2_l(q_l + p_l \cot \tau)}.$$

Put, for the moment,

$$\sum_l \beta^2_l(q_l + p_l \cot \tau) = f(E).$$

Then the quantity $(f(E))^{\frac{1}{2}}\ (E_\lambda - E)^{-\frac{1}{2}}\ \beta_k^{-1}\ \gamma_{\lambda k}$ may be regarded as the λ-component of the k^{th} member of a set of vectors, one for each split. But (*A*.9) states that each of these vectors has unit length, and that the scalar product of any pair of them is also one. Hence the vectors must all be identical, and, in fact

$$(A.10) \qquad \gamma^c_{\lambda k} = \gamma_\lambda \beta_k.$$

If the values of $\mathbf{R}^c$ given by (*A*.9) is used in (*A*.8) we obtain the form (3.16) for **U**. Thus (*A*.10) specializes Wigner's matrix to the case described by the compound nucleus hypothesis.

Moreover if we use (*A*.10) in (*A*.9) we find

$$(A.11) \qquad [\sum_k \beta^2_k(q_k + p_k \cot \tau)]^{-1} = \sum_\lambda \frac{\gamma^2_\lambda}{E_\lambda - E}.$$

Thus this function has simple poles at each E_λ, and cot $\tau(E_\tau) = 0$ when (cf. Equation 3.28)

$$(A.12) \qquad E_\tau = E_\lambda - \tfrac{1}{2}\Gamma_\lambda(\sum_j \beta^2_j q_j)(\sum_j \beta^2_j p_j)^{-1}.$$

This establishes the relation between the resonant energies E_λ in Wigner's theory, and the resonance condition cot $\tau = 0$ of the present work. The energy values E_λ are defined by a boundary condition imposed at the surface S on solutions of the Schroedinger equation in the interior region. It may be that these are not the only possible resonances. Since cot $\tau = 0$ gives a maximum value of cross sections, this condition should specify a resonance of any kind.

REFERENCES

1. Beiduck, F. M., Pruett, J. R., and Konopinski, E. J. Phys. Rev. 77: 622. 1950.
2. Bohr, N. Nature, 137: 344. 1936.
3. Feshbach, H., Peaslee, D. C., and Weisskopf, V. F. Phys. Rev. 71: 145. 1947.
4. Ghoshal, S. N. Phys. Rev. 80: 939. 1950.
5. Wigner, E. P. Gruppentheorie und ihre Anwendung auf die Quantenmechanik der Atomspektren. Braunschweig, 1931.
6. Wigner, E. P. and Eisenbud, L. Phys. Rev. 72: 29. 1947.
7. Wigner, E. P. Phys. Rev. 73: 1002. 1948.

AN ANALYSIS OF THE SELF-ENERGY PROBLEM FOR THE ELECTRON IN QUANTUM ELECTRODYNAMICS[1]

By P. N. Daykin[2]

ABSTRACT

Feynman's *S*-matrix for the self-energy of the free resting electron is evaluated without the restriction that the virtual photons in the intermediate state have only positive energy. Both the one-electron theory and the hole theory of the positron are treated. It is shown that in the one-electron theory the normally quadratically divergent transverse part of the self-energy vanishes if the photon field is assumed to be symmetric in positive and negative energies. A similar theorem does not hold in the hole theory. A particular type of interaction leads to a vanishing self-energy in one-electron theory. However, this does not solve the self-energy problem, as in this case radiation corrections to scattering would vanish as well. The *S*-matrix for the self-energy of a bound electron is evaluated in a similar manner. The decay probability for an excited state is calculated as the imaginary part of the self-energy. The correct value is obtained only in hole theory and in interaction with positive energy photons. In the special case in which the external field is a uniform magnetic field, again only hole theory with this same interaction gives the correct value for the anomalous magnetic moment.

I. INTRODUCTION

The self-energy problem of the electron can be briefly described as follows. According to the formalism of quantum mechanics a resting electron can emit and re-absorb virtual photons, thus spending part of its lifetime in intermediate states in which it has a kinetic energy. If the time average of this kinetic energy plus the energy of the virtual photons is computed, it turns out to be infinite, the divergence stemming from the contributions of the intermediate states with high energy. The divergence is quadratic in case of the ordinary one-electron theory, and is logarithmic if the hole theory of the positron is adopted (10, 11).

These well known calculations are based explicitly or implicitly on the assumption that only positive energy states are available to the virtual photons. However there is up to now no conclusive experimental evidence that this assumption is correct. Dirac (1) has suggested that a relativistic quantum theory of radiation should treat the positive and negative energies symmetrically. Pauli (7) has shown that this treatment leads to a vanishing transverse part of the self-energy in one-electron theory; but if the hole theory of the positron is adopted a similar theorem does not hold. Addition of effects of positive and negative energy photons in fact destroys the logarithmic character of the divergence, leaving quadratically divergent terms. Muto and Inoue (6) report that the divergence in the fourth order self-energy in one-electron theory is not eliminated by introducing negative energy quanta.

[1] *Manuscript received September 19, 1951.*
Contribution from the Department of Physics, The University of British Columbia, Vancouver, B.C.
[2] *Holder of a National Research Council of Canada Fellowship.*

It is the purpose of this paper to investigate the self-energy problem without the restriction that the photon field contains only positive frequencies. For this investigation use has been made of the simplified form of quantum electrodynamics described by Feynman (2, 3). It is shown that the distinction between one-electron theory and hole theory and between the different types of interaction can be made by defining different contours for evaluating Feynman's S-matrix for the free electron self-energy. Pauli's results are verified if the effects of Dirac's λ-limiting process are neglected. In addition it is found that there exists a particular type of interaction corresponding to a certain solution of the Maxwell equations discussed by Stueckelberg (9) which leads to a vanishing self-energy in one-electron theory. However, this does not represent a satisfactory solution of the self-energy problem, as in this case the radiation corrections to scattering would vanish as well. Thus it would be impossible in this case to obtain theoretically expressions for experimentally observed effects like the Lamb shift and the anomalous magnetic moment of the electron.

The S-matrix for the bound electron self-energy is evaluated for the same contours as used in the free electron case. The self-energy is in general not real; the negative of the imaginary part gives the decay probability per unit time for the initial bound state. It is found that the correct decay probability arises from only one contour, the contour which corresponds to the hole theory of the positron and interaction with only positive energy photons. In the special case where the external field is a homogeneous magnetic field this same contour is again the only one which gives the correct expression for the anomalous magnetic moment found by Schwinger (8) and Luttinger (5).

II. EVALUATION OF THE FREE ELECTRON SELF-ENERGY MATRIX

The notation introduced by Feynman will be used throughout this paper. According to Feynman (3) the self-energy of the free electron can be written as the single integral

$$\Delta E = \frac{e^2}{\pi i}\bar{u}\int \gamma_\mu(\mathbf{p} - \mathbf{k} - m)^{-1}\gamma_\mu k^{-2} d^4k u \tag{1}$$

where $d^4k = (2\pi)^{-2}dk_4 d^3\vec{K}$, $(\mathbf{p} - \mathbf{k} - m)^{-1} = (\mathbf{p} - \mathbf{k} + m)\{(p-k)^2 - m^2\}^{-1}$ and u is a solution of the Dirac equation.

$$(\mathbf{p} - m)u = 0,\ u^*u = 1,\ \bar{u} = u^*\beta. \tag{2}$$

The left hand side of (1) can therefore be written as $\Delta E u^* u$ or since

$$E\Delta E = m\Delta m \text{ and } u^*u = \bar{u}\frac{E}{m}u,$$

as $\Delta m \bar{u} u$. It is convenient to carry out the summation over μ in (1) right away by using

$$\begin{aligned} &\gamma_\mu \mathbf{A}\gamma_\mu = -2\mathbf{A} \text{ for any } \mathbf{A} = A_\mu\gamma_\mu \\ &\gamma_\mu\gamma_\mu = 4,\ \mathbf{p}u = mu \end{aligned} \tag{3}$$

thus removing the γ-matrices and obtain

$$\Delta m(\bar{u}u) = \frac{2e^2}{\pi i}\bar{u}(mJ_1 + \mathbf{J}_2)u \tag{4}$$

with

$$J_{1;2_\nu} = \int (1;k_\nu)\{(p-k)^2 - m^2\}^{-1}k^{-2}d^4k. \tag{5}$$

It is now the integration over k_4 which exhibits fully the advantage of treating the problem in momentum space, as the distinctions between one-electron theory and hole theory and the different kinds of interaction field can most conveniently be made by assigning appropriate conventions regarding the positions of the poles of the Integrals (5). As we are interested only in the case of the resting electron ($\vec{p} = 0$), we need calculate only J_1 and J_{24}, because $\vec{J}_2$ vanishes in this case by symmetry. If the integration over the variable k_4 is performed by completing the contour with a semicircle above the real k_4-axis the Integrals (5) can be written

$$J_{1;24} = 2i\int_0^\infty K^2 dK \sum R_{1;24} \tag{6}$$

where $\sum R_{1;24}$ has to be carried out over the respective residues at the poles positioned above the real k_4-axis. Each of the integrals (5) has four poles along the k_4-axis which are tabulated together with the corresponding residues of J_1 and J_{24} in Table I.

TABLE I

Poles	*Residues of* J_1	*Residues of* J_{24}
$k_4^{\mathrm{I}} = E - \sqrt{K^2+m^2}$	$R_1^{\mathrm{I}} = \dfrac{m^2 + E\sqrt{K^2+m^2}}{4m^2K^2\sqrt{K^2+m^2}}$	$R_{24}^{\mathrm{I}} = -\dfrac{E}{4m^2\sqrt{K^2+m^2}}$
$k_4^{\mathrm{II}} = -K$	$R_1^{\mathrm{II}} = -\dfrac{E}{4m^2K^2}$	$R_{24}^{\mathrm{II}} = +\dfrac{E}{4m^2K}$
$k_4^{\mathrm{III}} = +K$	$R_1^{\mathrm{III}} = -\dfrac{E}{4m^2K^2}$	$R_{24}^{\mathrm{III}} = -\dfrac{E}{4m^2K}$
$k_4^{\mathrm{IV}} = E + \sqrt{K^2+m^2}$	$R_1^{\mathrm{IV}} = -\dfrac{(m^2 - E\sqrt{K^2+m^2})}{4m^2K^2\sqrt{K^2+m^2}}$	$R_{24}^{\mathrm{IV}} = +\dfrac{E}{4m^2\sqrt{K^2+m^2}}$

In the computation of this table extensive use has been made of the relation

$$p_4^2 = E^2 = m^2$$

for the rest electron.

The distinction between one-electron theory and hole theory is now made in the usual way (2) by assigning for one-electron theory an infinitesimal positive imaginary part to E, thus bringing k_4^{I} and k_4^{IV} above the real k_4-axis, while assignment of an infinitesimal negative imaginary part to m leads to the hole theory, bringing k_4^{I} above and k_4^{IV} below the real k_4-axis. There remain then four possibilities for locating the poles k_4^{II} and k_4^{III}, each corresponding to a different kind of interaction field as listed in Table II.

TABLE II

Position of poles k_4^{II} and k_4^{III} relative to the real k_4-axis	*Corresponding photon kernel in ordinary space*	*Type of interaction field*
k_4^{II} *above* k_4^{III} *below*	$\delta_+(S_{21}^2) = \frac{1}{2r}\{\delta_+(t_{21} - r_{21}) + \delta_+(-t_{21} - r_{21})\}$	*Positive energy photon field*
k_4^{II} *below* k_4^{III} *above*	$\delta_-(S_{21}^2) = \frac{1}{2r}\{\delta_-(t_{21} - r_{21}) + \delta_-(-t_{21} - r_{21})\}$	*Negative energy photon field*
k_4^{II} *below* k_4^{III} *below*	$D_+(x_{21}) = \begin{cases} \frac{1}{r}\delta(t_{21} - r_{21}); t_2 > t_1 \\ 0 \quad ; t_2 < t_1 \end{cases}$	*Stueckelberg field*
k_4^{II} *above* k_4^{III} *above*	$D_-(x_{21}) = \begin{cases} 0 \quad ; t_2 > t_1 \\ \frac{1}{r}\delta(t_{21} + r_{21}); t_2 < t_1 \end{cases}$	*Stueckelberg field*

The photon kernels* of main interest from the point of view of this paper are the δ_+ and δ_- functions. It is clear from the positions of the poles that the δ_+ and δ_- functions differ only in the sign of $K = (\vec{K}\cdot\vec{K})^{\frac{1}{2}}$. The interactions of the Stueckelberg type are included for the sake of completeness.

According to a remark by Feynman (7), the δ_+ or δ_- interaction between two electrons leads to a theory which, in the classical limit, exhibits interactions respectively through retarded or advanced potentials. Thus in the self-energy problem the motion of a wave packet under the δ_+ or δ_- self-interaction would lead, in the classical limit, to the motion of an electron in interaction with its proper retarded or advanced field.

Thus the integration over k_4 in (5) as outlined above yields the following complete self-energy table (III) for the free resting electron.

**The term "photon kernel" was introduced in analogy with the term "electron kernel" to mean any singular solution of the Maxwell equations for the potentials.*

TABLE III

Photon kernel	$\sum R$	Δm
One-electron theory		
δ_+	$R^{\mathrm{I}} + R^{\mathrm{II}} + R^{\mathrm{IV}}$	$\frac{e^2}{\pi m}\left\{E\int dK + \int K dK\right\}$
δ_-	$R^{\mathrm{I}} + R^{\mathrm{III}} + R^{\mathrm{IV}}$	$\frac{e^2}{\pi m}\left\{E\int dK - \int K dK\right\}$
D_+	$R^{\mathrm{I}} + R^{\mathrm{IV}}$	$\frac{2e^2}{\pi m}E\int dK$
D_-	$R^{\mathrm{I}} + R^{\mathrm{II}} + R^{\mathrm{III}} + R^{\mathrm{IV}}$	*Zero*
Hole theory		
δ_+	$R^{\mathrm{I}} + R^{\mathrm{II}}$	$\frac{e^2 m}{\pi}\int dK\left\{\frac{1}{\sqrt{K^2+m^2}} - \frac{K^2}{m^2\sqrt{K^2+m^2}} + \frac{K}{m^2}\right\}$
δ_-	$R^{\mathrm{I}} + R^{\mathrm{III}}$	$\frac{e^2 m}{\pi}\int dK\left\{\frac{1}{\sqrt{K^2+m^2}} - \frac{K^2}{m^2\sqrt{K^2+m^2}} - \frac{K}{m^2}\right\}$
D_+	R^{I}	$\frac{e^2 m}{\pi}\int dK\left\{\frac{1}{\sqrt{K^2+m^2}} + \frac{E}{m^2} - \frac{K^2}{m^2\sqrt{K^2+m^2}}\right\}$
D_-	$R^{\mathrm{I}} + R^{\mathrm{II}} + R^{\mathrm{III}}$	$\frac{e^2 m}{\pi}\int dK\left\{\frac{1}{\sqrt{K^2+m^2}} - \frac{E}{m^2} - \frac{K^2}{m^2\sqrt{K^2+m^2}}\right\}$

It can be seen that in the one-electron theory the assumption that the interaction is $\frac{1}{2}(\delta_+ + \delta_-)$, which corresponds to Dirac's assumption that the photon field is symmetric in positive and negative energy states, leads to cancellation of the quadratically divergent transverse part of the self-energy. It is interesting to note that interaction of the Stueckelberg type, D_-, leads to vanishing self-energy in one-electron theory. Unfortunately this does not solve the self-energy problem, as in this case radiation corrections to scattering characterized by integrals of the type

$$\frac{e^4}{\pi i}\int \gamma_\mu(\mathbf{p}_2 - \mathbf{k} - m)^{-1}\mathbf{a}(\mathbf{p}_1 - \mathbf{k} - m)^{-1}\gamma_\mu k^{-2} d^4k$$

will vanish as well, because all poles of such integrals too are then located above the real k_4-axis. This would contradict experimental evidence like the Lamb shift and the anomalous magnetic moment of the electron. In the hole theory

it is only the δ_+ interaction which makes the divergence logarithmic, all other interactions leading to quadratically divergent expressions for Δm. Thus a linear combination of the form $\frac{1}{2}(\delta_+ + \delta_-)$ interaction would remove the term k/m^2 in the integrand of Δm and destroy the logarithmic character of the divergence leaving a quadratically divergent expression for Δm.

III. SELF-ENERGY MATRIX FOR AN ELECTRON IN AN EXTERNAL FIELD

Feynman's S-matrix for the free electron self-energy can be generalized for an electron in an external field, $\mathbf{A}$, by replacing the free electron kernel $K_+(2,1)$ by $K_+^{(A)}(2,1)$, the kernel in hole theory for the Dirac equation

$$(i\nabla - \mathbf{A} - m)K_+^{(A)}(2,1) = i\delta(2,1) \tag{7}$$

which may be written as the sum over eigenfunctions of Equation (7)

$$K_+^{(A)}(2,1) = \begin{cases} \sum_+ \phi_n(2)\bar{\phi}_n(1)e^{-iE_n t_{21}} & t_2 > t_1 \\ \sum_- \phi_n(2)\bar{\phi}_n(1)e^{-iE_n t_{21}} & t_2 < t_1 . \end{cases}$$

The plane wave functions appearing in the S-matrix must correspondingly be replaced by eigenfunctions of Equation (7). The diagonal element of the S-matrix for the state m then becomes

$$S_{mm} = -ie^2\int_0^T dt_1\int_{-\infty}^{\infty} dt_2\iint \bar{\phi}_m(2)\gamma_\mu K_+^{(A)}(2,1)\gamma_\mu\phi_m(1) \cdot e^{iE_m(t_2-t_1)}\delta_+(S_{21}^2)d^3x_2 d^3x_1$$

where

$$\delta_+(S_{21}^2) = -\frac{1}{\pi}\int e^{-ik\cdot x_{21}}\mathbf{k}^{-2}d^4k.$$

Clearly S is proportional to the integration time T so we can write for the self-energy

$$W_m = e^2\int_{-\infty}^{\infty} dt_2\iint \bar{\phi}_m(2)\gamma_\mu K_+^{(A)}(2,1)\gamma_\mu\phi_m(1)\, e^{iE_m(t_2-t_1)}\delta_+(S_{21}^2)d^3x_2 d^3x_1.$$

In one-electron theory $K_+^{(A)}$ is replaced by the one-electron kernel, $K_0^{(A)}$, given by

$$K_0^{(A)}(2,1) = \begin{cases} \sum_n \phi_n(2)\bar{\phi}_n(1)e^{-iE_n t_{21}}; & t_2 > t_1 \\ 0 & ; t_2 < t_1. \end{cases}$$

The distinction between the different types of interaction is now made as in Section II by assigning different conventions for evaluating the poles of k^{-2} in the k_4 complex plane. The necessary splitting of the t_2 integration into two ranges $(-\infty, t_1)$ and (t_1, ∞), gives rise to an imaginary part to the self-energy according to the relation

$$\frac{i}{\pi}\int_0^\infty e^{\mp iwt}dt = i\delta_\pm(w) = i\delta(w) \pm \frac{1}{\pi w}.$$

The integration over $\vec{K}$ for the imaginary part can most conveniently be completed by making the dipole approximation and using the formula

$$\int \frac{d^3K}{K}\delta(E-K) = \begin{cases} 4\pi E; E > 0 \\ 0 \quad ; E \leqslant 0. \end{cases}$$

If we write

$$W_m = \Delta E_m - i\frac{\gamma_m}{2}$$

then, in $|\exp(-iW_m t)|^2$, γ_m produces an exponentially decreasing amplitude with time; so γ_m is just the decay probability per unit time for the state m. Finally the relations $\bar{\phi} = \phi^*\beta$ and $\gamma_\mu = \beta a_\mu$ were used to simplify the expressions for ΔE_m and γ_m which are listed in Columns 2 and 3 of Table IV.

An inspection of Column 3 shows that only the δ_+ interaction in the hole theory gives the correct rate of radiation from an excited state, as shown below. For simplicity the table is restricted to the case in which E_m is positive. In the one-electron theory the δ_+ interaction allows the electron to fall into the negative energy states with emission of a positive energy photon and the δ_- interaction allows transitions upward with emission of a negative energy photon. The D_+ interaction gives the same result as the $(\delta_+ + \delta_-)$ interaction. In the hole theory the δ_- interaction contains a second set of terms allowing transitions to negative energy states. These terms must be associated with the possibility of pair creation out of the vacuum along with emission of a negative energy photon, but they will not be discussed further here. The D_+ and D_- functions allow transitions both up and down to occur.

It appears that the longitudinal and scalar photons as well as the transverse photons contribute to γ, but the first two contributions exactly cancel. For proof, the matrix element V^μ_{mn} can be written in momentum space as

$$\int \phi^*_m(P-K)\phi_n(P)(\bar{u}_m(P-K)\gamma_\mu U_n(P))d^3P$$

in which $u_n(P)$ is the Dirac spinor and $\phi_n(P)$ is the coefficient in the expansion of $\phi_n(x)$ in momentum eigenfunctions. It follows from the Dirac equation $\mathbf{p}u(P) = mu(P)$ and its conjugate that

$$k_\mu V^\mu_{mn} = 0.$$

For a free photon $k_4 = K$ so

$$V^4_{mn} = V^k_{mn}.$$

Thus

$$-|V^\mu_{mn}|^2 = |\vec{V}_{mn}|^2 - |V^k_{mn}|^2,$$

which is just the sum over the transverse photons. Finally, averaging over all directions $\vec{K}$ and making the usual nonrelativistic approximations lead to the radiation formula given by Heitler (4). For a second proof see Feynman (2).

TABLE IV

Photon kernel	ΔE_m	γ_m
One-electron theory		
δ_+	$\frac{e^2}{4\pi^2}\int\frac{d^3K}{K}\sum_{E_n}\frac{\lvert V\rvert^2}{E_n-E_m+K}$	$-2e^2\sum_{E_n<E_m}\lvert V\rvert^2(E_m-E_n)$
δ_-	$-\frac{e^2}{4\pi^2}\int\frac{d^3K}{K}\sum_{E_n}\frac{\lvert V\rvert^2}{E_n-E_m-K}$	$2e^2\sum_{0<E_m<E_n}\lvert V\rvert^2(E_n-E_m)$
D_+	$\frac{e^2}{4\pi^2}\int\frac{d^3K}{K}\sum_{E_n}\left\{\frac{\lvert V\rvert^2}{E_n-E_m+K}-\frac{\lvert V\rvert^2}{E_n-E_m-K}\right\}$	$2e^2\{\sum_{0<E_m<E_n}\lvert V\rvert^2(E_n-E_m)-\sum_{E_n<E_m}\lvert V\rvert^2(E_m-E_n)\}$
D_-	*Zero*	*Zero*
Hole theory		
δ_+	$\frac{e^2}{4\pi^2}\int\frac{d^3K}{K}\left\{\sum_+\frac{\lvert V\rvert^2}{E_n-E_m+K}+\sum_-\frac{\lvert V\rvert^2}{E_n-E_m-K}\right.$	$-2e^2\sum_{0<E_n<E_m}\lvert V\rvert^2(E_m-E_n)$
δ_-	$-\frac{e^2}{4\pi^2}\int\frac{d^3K}{K}\left\{\sum_+\frac{\lvert V\rvert^2}{E_n-E_m-K}+\sum_-\frac{\lvert V\rvert^2}{E_n-E_m+K}\right\}$	$2e^2\{\sum_{0<E_m<E_n}\lvert V\rvert^2(E_n-E_m)-\sum_{E_n<0<E_m}\lvert V\rvert^2(E_m-E_n)\}$
D_+	$\frac{e^2}{4\pi^2}\int\frac{d^3K}{K}\sum_+\left\{\frac{\lvert V\rvert^2}{E_n-E_m+K}-\frac{\lvert V\rvert^2}{E_n-E_m-K}\right\}$	$2e^2\sum_+\lvert V\rvert^2(E_m-E_n)$
D_-	$-\frac{e^2}{4\pi^2}\int\frac{d^3K}{K}\sum_-\left\{\frac{\lvert V\rvert^2}{E_n-E_m+K}-\frac{\lvert V\rvert^2}{E_n-E_m-K}\right\}$	$2e^2\sum_-\lvert V\rvert^2(E_m-E_n)$

$$V=V^{\mu}_{mn}=\int\phi^*_m(x)a_\mu e^{-ikr}\phi_n(x)d^3x,\quad a_\mu a_\mu=1-a,a$$

IV. THE ANOMALOUS MAGNETIC MOMENT OF THE ELECTRON

Luttinger has shown that the anomalous magnetic moment may be obtained from the self-energy of the electron in a uniform magnetic field by taking advantage of the existence of a particular unperturbed state for which $E = m$. For this state the contribution to ΔE due to the change in mass of the electron is just Δm so that the mass renormalization requires only the direct subtraction of Δm. We have followed closely the method of Luttinger in calculating the anomalous magnetic moment predicted by the self-energy formulas given in Column 2. However the factor $|V|^2$ contains the sum over all four polarization directions, μ, thus including the Coulomb interaction directly as the effect of interaction with longitudinal photons. The details of the expansion of the expression for ΔE in powers of the magnetic field strength, H, can be found in Luttinger's paper. The term independent of H in each case is just the corresponding expression for Δm from Table III.

It was found that the correct expression for the anomalous magnetic moment given by Luttinger and Schwinger is obtained only from the δ_+ interaction in hole theory. The remaining six nonvanishing expressions lead to a quadratically divergent expression for the moment, even after subtraction of Δm.

V. CONCLUSION

It must be concluded that there is no solution to the self-energy problem possible by introducing negative energy photons without some modifications.

VI. ACKNOWLEDGEMENTS

I should like to thank the National Research Council of Canada for support of this work by the award of a Fellowship and Prof. F. A. Kaempffer for suggesting the topic and for many valuable discussions.

REFERENCES

1. DIRAC, P. A. M. Proc. Roy. Soc. A, 180: 1. 1942.
2. FEYNMAN, R. P. Phys. Rev. 76: 749. 1949.
3. FEYNMAN, R. P. Phys. Rev. 76: 769. 1949.
4. HEITLER, W. Quantum theory of radiation. Oxford University Press. p. 106.
5. LUTTINGER, J. M. Phys. Rev. 74: 893. 1948.
6. MUTO, T. and INOUE, K. Progress Theor. Phys. 5: 1033. 1950.
7. PAULI, W. Revs. Modern Phys. 15: 175. 1943.
8. SCHWINGER, J. Phys. Rev. 73: 416. 1948.
9. STUECKELBERG, C. Helv. Phys. Acta, 11: 225. 1938.
10. WEISSKOPF, V. Z. Physik, 89: 27. 1934.
11. WEISSKOPF, V. Phys. Rev. 56: 72. 1939.

NOTE

A New Band System Excited in Acetylene*

During a search for emission spectra of polyatomic free radicals a new band system was observed in a high frequency discharge through fast flowing acetylene. Well exposed plates on a medium quartz instrument have been obtained which show the system to be very complex. It has unfortunately not been possible to excite the system with sufficient intensity to allow examination under higher dispersion.

The conditions in the discharge were chosen so that excitation of polyatomic rather than diatomic species occurred. Vapors were pumped rapidly through a discharge tube 4 ft. long and 4 in. in diameter at a pressure of about 0.1 mm. The tube was placed in a coil which was loosely coupled to the oscillator coil of a 5 kw. 30 m. oscillator, the power input being only a small fraction of the possible maximum.

The conditions are similar to those obtained by Schüler in a specially designed glow discharge (4). The spectrum obtained with the present type of discharge through formaldehyde showed the 'fluorescence' bands of formaldehyde and some of the ethylene flame bands and duplicates that obtained by Schüler (3). Formic acid gave a band system, first discovered by Terenin and Neujmin in fluorescence in formic acid and subsequently studied by Dyne and Style. This is due to a polyatomic radical (2, 6).

The band system excited in acetylene consists of about 40 bands lying between λ2400 and λ2900 superposed on a continuum which is strong at the long wave length end of the system, but weak compared with the bands at λ2500 (Fig. 1). In addition to this system, bands due to C_2 (Swan, Deslandres-d'Azambuja, Mulliken) and CH (λ4300, λ3900, λ3143) are prominent. The "λ4050" group (1) appears weakly. A few of the 5B and fourth positive bands of CO overlap the new system, but these are weak and easily distinguished.

Wave lengths of the more prominent features are given in Table I. Bands in the region λ2400–λ2600 are generally sharp and could be measured accurately. These are quoted to 0.1 Å. At longer wave lengths the intensity of the continuous background and the more diffuse nature of the bands made accurate measurements impossible, and their wave lengths are quoted to the nearest angstrom unit. The structure of the system is complex; some weak "lines" and "heads" and signs of structure are present which are not recorded in the table. Many of the strong bands are double-headed, the separation of the heads being about one angstrom unit and degraded to the red.

The band system could be due to acetylene or to some radical derived from it (e.g. C_2H). This system may be compared to the known absorption system (5) of acetylene which extends from shorter wave lengths to about λ2400, i.e. to

*Issued as N.R.C. No. 2633.

TABLE I

λ	Intensity	Description	λ	Intensity	Description
2399.3	1		2573.9	1	d.
2406.2	1	d.h.	2580.4	3	d.
2410.5	1		2598.2	2	
2419.9	3	d.h.	2609	1	
2423.0	2	l.	2615	2	d.
2424.0	1	l.	2618	2	l.
2425.1	1	l.	2629	4	b.
2430.0	1	l.	2642	3	d.h.
2441.1	3	d.h.	2693	3	b.
2454.6	1	s.	2698	1	
2461.3	3	d.h.	2716	1	
2466.3	3	l.	2723	4	d.h.
2478.9	3	d.h.	2769	2	b.
2493.5	3		2787	2	
2499.4	4		2869	3	b.
2514.5	2	d.h.	2920	1	
2522.1	4	d.h.			
2526.7	3	d.			
2541.0	2				
2559.3	2	d.h.			
2565.6	4	l.			

Note: s. = sharp, b. = broad, d. = diffuse, l. = linelike, d.h. = double-headed.

the short wave length limit of the emission system. The absorption system is very complex; the bands are double-headed (separation about one angstrom unit) and are degraded to the red. Progressions of about 1050 cm.$^{-1}$ are observed. These are all points of similarity with the emission system in which short progressions of ~1030 cm.$^{-1}$ can be picked out. In addition spacings of ~1900 cm.$^{-1}$, the characteristic $C \equiv C$ frequency, can be picked out.

It is tempting to suppose that the emission system is a long wave length extension of the absorption system, but the present evidence does not allow any decisive decision to be made.

The author wishes to acknowledge the help and advice given during the course of this work by Dr. G. Herzberg, F.R.S., and Dr. A. E. Douglas.

1. Douglas, A. E. Astrophys. J. 114: 466. 1951.
2. Dyne, P. J. and Style, D. W. G. Discussions of the Faraday Society, 2: 159. 1942, and unpublished work.
3. Schüler, H. and Reinebeck, L. Z. Naturforsch. A, 5: 604. 1950.
4. Schüler, H. and Woeldike, A. Chem. Tech. (Berlin), 15: 99. 1942.
5. Sho-Chow Woo, Ta-Kong Liu, T. C. Chu, and Wu Chih. J. Chem. Phys. 6: 240. 1938.
6. Terenin, A. and Neujmin, H. Acta Physicochim. U.R.S.S. 5: 465. 1936.

Received August 21, 1951.
P. J. Dyne[1]
Division of Physics,
National Research Council,
Ottawa, Canada.

[1]*National Research Council of Canada Postdoctorate Fellow.*

(Note: *Plate 1 (Fig. 1) follows.*)

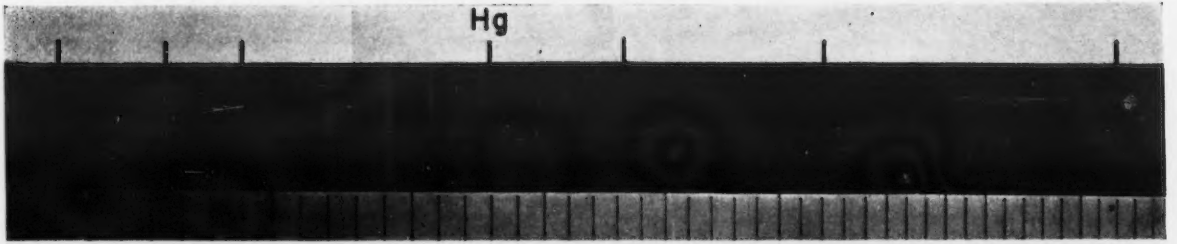

FIG. 1. Band system excited in acetylene taken on a medium quartz spectrograph. Exposure 1½ hr. Eastman 103*a*-0 plate. Bands marked with a leading line are assigned to CO. The wave length scale is slightly in error; the indicated wave lengths are about 4 Å. too low.

CANADIAN JOURNAL OF PHYSICS

Notice to Contributors

GENERAL: Manuscripts should be typewritten, double spaced, and the **original and one extra copy** submitted. Style, arrangement, spelling, and abbreviations should conform to the usage of this Journal. Names of all simple compounds, rather than their formulas, should be used in the text. Greek letters or unusual signs should be written plainly or explained by marginal notes. Superscripts and subscripts must be legible and carefully placed. Manuscripts should be carefully checked before being submitted, to reduce the need for changes after the type has been set. If authors require changes to be made after the type is set, they will be charged for changes that are considered to be excessive. **All pages, whether text, figures, or tables, should be numbered.**

ABSTRACT: **An abstract of not more than about 200 words, indicating the scope of the work and the principal findings, is required.**

ILLUSTRATIONS:

(i) *Line Drawings:* All lines should be of sufficient thickness to reproduce well. Drawings should be carefully made with India ink on white drawing paper, blue tracing linen, or co-ordinate paper **ruled in blue only**; any co-ordinate lines that are to appear in the reproduction should be ruled in black ink. Paper ruled in **green, yellow, or red should not be used** unless it is desired to have all the co-ordinate lines show. Lettering and numerals should be neatly done in India ink preferably with a stencil (**do not use typewriting**) and be of such size that they will be legible and not less than one millimeter in height when reproduced in a cut three inches wide. All experimental points should be carefully drawn with instruments. Illustrations need not be more than two or three times the size of the desired reproduction, but the ratio of height to width should conform with that of the type page. **The original drawings and one set of small but clear photographic copies are to be submitted.**

(ii) *Photographs:* Prints should be made on glossy paper, with strong contrasts; they should be trimmed to remove all extraneous material so that essential features only are shown. Photographs should be submitted **in duplicate**; if they are to be reproduced in groups, one set should be so arranged and mounted on cardboard with rubber cement; the duplicate set should be unmounted.

(iii) *General:* **The author's name, title of paper, and figure number should be written in the lower left hand corner (outside the illustration proper) of the sheets on which the illustrations appear.** Captions should not be written on the illustrations, but typed together at the end of the manuscript. All figures (including each figure of the plates) should be numbered consecutively from 1 up (arabic numerals). **Each figure should be referred to in the text.** If authors desire to alter a cut, they will be charged for the new cut.

TABLES: Each table should be typed on a separate sheet. Titles should be given for all tables, which should be numbered in Roman numerals. Column heads should be brief and textual matter in tables confined to a minimum. **Each table should be referred to in the text.**

REFERENCES: **These should be listed alphabetically by authors' names, numbered in that order, and placed at the end of the paper.** The form of literature citation should be that used in this Journal. **Titles of papers should not be given.** The first page only of the references cited should be given. **All citations should be checked with the original articles.** Each citation should be referred to in the text by means of the key number.

REPRINTS: A total of 50 reprints of each paper, without covers, are supplied free to the authors. Additional reprints will be supplied according to a prescribed schedule of charges. On request, covers can be supplied at cost.

Approximate charges for reprints may be calculated from the number of printed pages, obtained by multiplying by 0.6 the number of manuscript pages (double-spaced typewritten sheets, 8½ in. by 11 in.) and making allowance for space occupied by line drawings and half-tones (not inserts). The cost per page is tabulated at the back of the reprint request form sent with the galley.

Contents

Printed in Canada *by* University of Toronto Press